

GenCore version 5.1.6
Copyright (c) 1993 - 2004 Compugen Ltd.

OM protein - protein search, using sw model

Run on: October 27, 2004, 08:33:39 ; Search time 132 Seconds
(without alignments)
827.733 Million cell updates/sec

Title: US-09-867-570-2

Perfect score: 1763

Sequence: 1 MESKSSWVIRLGFSLSDSTI.....EGCGWLPQETLISGRLQ 337

Scoring table: BLOSUM62

Searched: 1370721 seqs, 324215800 residues

Total number of hits satisfying chosen parameters: 1370721

Minimum DB seq length: 0

Maximum DB seq length: 200000000

Post-processing: Maximum Match 0%

Listing first 45 summaries

Database :
1: /cgn2_6/ptodata/1/pubppa/US07_PUBCOMB.pep:*
2: /cgn2_6/ptodata/1/pubppa/US06_PUBCOMB.pep:*
3: /cgn2_6/ptodata/1/pubppa/US05_PUBCOMB.pep:*
4: /cgn2_6/ptodata/1/pubppa/US04_PUBCOMB.pep:*
5: /cgn2_6/ptodata/1/pubppa/US03_PUBCOMB.pep:*
6: /cgn2_6/ptodata/1/pubppa/US02_PUBCOMB.pep:*
7: /cgn2_6/ptodata/1/pubppa/US01_PUBCOMB.pep:*
8: /cgn2_6/ptodata/1/pubppa/US00_PUBCOMB.pep:*
9: /cgn2_6/ptodata/1/pubppa/US09_PUBCOMB.pep:*
10: /cgn2_6/ptodata/1/pubppa/US08_PUBCOMB.pep:*
11: /cgn2_6/ptodata/1/pubppa/US07_PUBCOMB.pep:*
12: /cgn2_6/ptodata/1/pubppa/US06_PUBCOMB.pep:*
13: /cgn2_6/ptodata/1/pubppa/US05_PUBCOMB.pep:*
14: /cgn2_6/ptodata/1/pubppa/US04_PUBCOMB.pep:*
15: /cgn2_6/ptodata/1/pubppa/US03_PUBCOMB.pep:*
16: /cgn2_6/ptodata/1/pubppa/US02_PUBCOMB.pep:*
17: /cgn2_6/ptodata/1/pubppa/US01_PUBCOMB.pep:*
18: /cgn2_6/ptodata/1/pubppa/US00_PUBCOMB.pep:*
19: /cgn2_6/ptodata/1/pubppa/US09_PUBCOMB.pep:*
20: /cgn2_6/ptodata/1/pubppa/US08_PUBCOMB.pep:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	1763	100.0	337	11	US-09-867-570-2
2	1688	95.7	332	9	US-09-995-225-20
3	1688	95.7	332	10	US-09-995-225-20
4	1688	95.7	332	14	US-10-183-116-31
5	1688	95.7	332	14	US-10-825-567A-674
6	1688	95.7	332	15	US-10-072-012-529
7	1688	95.7	332	15	US-10-072-012-534
8	1688	95.5	332	14	US-10-401-397A-2
9	1688	95.5	332	14	US-10-292-798-1274
10	1683	95.5	332	15	US-10-072-012-530
11	1683	94.3	332	15	US-10-072-012-535
12	1663	93.0	314	15	US-10-391-074-2
13	1639	93.0	314	14	US-10-219-834-79

14	1557	88.3	302	14	US-10-237-467-10	Sequence 10, Appl
15	1415.5	80.3	1589	15	US-10-072-012-528	Sequence 528, App
16	1415.5	80.3	1589	15	US-10-072-012-532	Sequence 532, App
17	1383	78.4	322	14	US-10-183-116-16	Sequence 16, Appl
18	1383	78.4	322	14	US-10-079-364-4	Sequence 4, Appl
19	1383	78.4	322	14	US-10-017-161-1056	Sequence 1056, Ap
20	1383	78.4	322	14	US-10-240-998-4	Sequence 4, Appl
21	1383	78.4	322	14	US-10-321-807-20	Sequence 20, Appl
22	1383	78.4	322	14	US-10-237-467-12	Sequence 12, Appl
23	1383	78.4	322	14	US-10-292-798-898	Sequence 898, App
24	1383	78.4	322	15	US-10-016-248-81	Sequence 81, Appl
25	1383	78.4	322	15	US-10-072-012-172	Sequence 172, App
26	1383	78.4	322	15	US-10-072-012-527	Sequence 527, App
27	1383	78.4	322	15	US-10-072-012-531	Sequence 531, App
28	1383	78.4	322	15	US-10-343-650A-44	Sequence 44, Appl
29	1383	78.4	322	16	US-10-321-807-20	Sequence 20, Appl
30	1383	78.4	322	16	US-10-314-048A-20	Sequence 20, Appl
31	1367	77.5	322	9	US-09-995-225-18	Sequence 18, Appl
32	1367	77.5	322	10	US-09-995-225-18	Sequence 18, Appl
33	1367	77.5	322	14	US-10-183-116-33	Sequence 33, Appl
34	1367	77.5	322	14	US-10-225-567A-689	Sequence 689, App
35	1367	77.5	322	14	US-10-237-467-4	Sequence 4, Appl
36	1367	77.5	322	15	US-10-016-248-84	Sequence 84, Appl
37	1367	77.5	322	15	US-10-072-012-531	Sequence 531, App
38	1367	77.5	322	15	US-10-072-012-536	Sequence 536, App
39	1365	77.2	319	15	US-10-072-012-174	Sequence 174, App
40	1361	77.2	322	14	US-10-292-798-1042	Sequence 1042, Ap
41	1313.5	74.5	323	15	US-10-072-012-178	Sequence 178, App
42	1263	71.6	328	14	US-10-219-834-20	Sequence 20, Appl
43	1031.5	58.5	331	16	US-10-467-616-2	Sequence 2, Appl
44	989.5	56.1	330	9	US-09-826-508-12	Sequence 12, Appl
45	989.5	56.1	330	9	US-09-750-373-24	Sequence 24, Appl

ALIGNMENTS

RESULT 1
US-09-867-570-2
; Sequence 2, Application US/09867570
; Publication No. US20040076951A1
; GENERAL INFORMATION:
; APPLICANT: WEI, Ming-Hui et al.
; TITLE OF INVENTION: ISOLATED HUMAN G-PROTEIN COUPLED RECEPTORS, NUCLEIC ACID MOLECULES ENCODING HUMAN GPCR
; TITLE OF INVENTION: PROTEINS, AND USBS THEREOF
; FILE REFERENCE: CL000900-CIP
; CURRENT APPLICATION NUMBER: US/09/867,570
; CURRENT FILING DATE: 2001-05-31
; PRIOR APPLICATION NUMBER: 09/695,045
; PRIOR FILING DATE: 2000-10-25
; NUMBER OF SEQ ID NOS: 4
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 2
; LENGTH: 337
; TYPE: PRT
; ORGANISM: Human
US-09-867-570-2

Query Match 100.0%; Score 1763; DB 11; Length 337;
Best Local Similarity 100.0%; Pred. No. 7.4e-152;
Matches 337, Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 1 MESKSSWVIRLGFSLSDSTIPVIGTELTPINGREPTCYKOTISFPTGLCTIVSLVALTGN 60
DB 1 MESKSSWVIRLGFSLSDSTIPVIGTELTPINGREPTCYKOTISFPTGLCTIVSLVALTGN 60
QY 61 AVLTWLTGMRNRNAYSIYILNVADPLFLSGHIIICSPRLINIRHPISKIISPWTTP 120
DB 61 AVLTWLTGMRNRNAYSIYILNVADPLFLSGHIIICSPRLINIRHPISKIISPWTTP 120
QY 121 YPTGLSMLSAISTERCLSTLWPIWYHCRPRRYLSYVMCYLWALSILRSITLWMPCDFLF 180
DB 121 YPTGLSMLSAISTERCLSTLWPIWYHCRPRRYLSYVMCYLWALSILRSITLWMPCDFLF 180

Db 121 YFGLSMISAISTERCSTILMPWYHCRPRYLSSVMCVLLMALSLRSILEMFCDFLF 180
QY 181 SGADSWCETSDFTTIAVLVFLCVVLGSSSLVLRILCGSRKMPLTRLYTTLTVLVF 240
Db 181 SGADSWCETSDFTTIAVLVFLCVVLGSSSLVLRILCGSRKMPLTRLYTTLTVLVF 240
QY 241 LCGSPGICQWALFSRHLDMKVLFCVHLVSIPLSALNSSANPITTFVGSFRORONRQ 300
Db 241 LCGSPGICQWALFSRHLDMKVLFCVHLVSIPLSALNSSANPITTFVGSFRORONRQ 300
QY 301 NLKVLQRALQDTPREVDEGGWMLPQETLESGSRLEQ 337
Db 301 NLKVLQRALQDTPREVDEGGWMLPQETLESGSRLEQ 337

RESULT 2

US-09-995-225-20
Sequence 20, Application US/09995225
Publication No. US20020193584A1
GENERAL INFORMATION:
APPLICANT: Chen, Ruoping
APPLICANT: Dang, Zhong Liang
APPLICANT: Dang, Huong T.
APPLICANT: Lowitz, Kevin P.
APPLICANT: Pride, Cameron
TITLE OF INVENTION: Endogenous And No. US20020193584A1-Endogenous Versions of Human G
FILE REFERENCE: AREN-0308
CURRENT APPLICATION NUMBER: US/09/995,225
CURRENT FILING DATE: 2001-11-26
PRIOR APPLICATION NUMBER: 09/170,496
PRIOR FILING DATE: 1998-10-13
PRIOR APPLICATION NUMBER: PCT/US99/23938
PRIOR FILING DATE: 1998-10-13
PRIOR APPLICATION NUMBER: 60/253,404
PRIOR FILING DATE: 2000-11-27
PRIOR APPLICATION NUMBER: 60/255,366
PRIOR FILING DATE: 2000-12-12
PRIOR APPLICATION NUMBER: 60/270,286
PRIOR FILING DATE: 2001-02-20
PRIOR APPLICATION NUMBER: 60/282,365
PRIOR FILING DATE: 2001-04-06
PRIOR APPLICATION NUMBER: 60/270,266
PRIOR FILING DATE: 2001-02-20
PRIOR APPLICATION NUMBER: 60/282,032
PRIOR FILING DATE: 2001-04-06
PRIOR APPLICATION NUMBER: 60/282,358
PRIOR FILING DATE: 2001-04-06
PRIOR APPLICATION NUMBER: 60/282,356
PRIOR FILING DATE: 2001-04-06
PRIOR APPLICATION NUMBER: 60/290,917
PRIOR FILING DATE: 2001-05-14
PRIOR APPLICATION NUMBER: 60/309,208
PRIOR FILING DATE: 2001-07-31
NUMBER OF SEQ ID NOS: 67
SOFTWARE: PatentIn version 3.1
SEQ ID NO 20
LENGTH: 322
TYPE: PRT
ORGANISM: Artificial Sequence
FEATURE:
OTHER INFORMATION: No. US20020193584A1 Sequence
US-09-995-225-20

Query Match 95.7%; Score 1688; DB 9; Length 322;
Best Local Similarity 100.0%; Pred. No. 4.6e-145;
Matches 322; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 16 MDSTIPVAGTEPLINGREPTCYKOTLSFTGLCTIVLALVAGAAVVMILGCMRRNA 75
Db 1 MDSTIPVAGTEPLINGREPTCYKOTLSFTGLCTIVLALVAGAAVVMILGCMRRNA 60
QY 76 VSIYIILNVADFLFSGHIIICSPRLINIRHPISKILSPWTFPYTGLSMLSAISTER 135

Db 61 VSIYIILNVADFLFSGHIIICSPRLINIRHPISKILSPWTFPYTGLSMLSAISTER 120
QY 136 CLSTIMPWYHCRPRYLSSVMCVLLMALSLRSILEMFCDFLFSGADSWCETSDFT 195
Db 121 CLSTIMPWYHCRPRYLSSVMCVLLMALSLRSILEMFCDFLFSGADSWCETSDFT 180
QY 196 IAMLVFLCVVLGSSSLVLRILCGSRKMPLTRLYTTLTVLVFLGSPGICQWALFS 255
Db 181 IAMLVFLCVVLGSSSLVLRILCGSRKMPLTRLYTTLTVLVFLGSPGICQWALFS 240
QY 256 RHLDMKVLFCVHLVSIPLSALNSSANPITTFVGSFRORONRQNLKVLQRALQDTP 315
Db 241 RHLDMKVLFCVHLVSIPLSALNSSANPITTFVGSFRORONRQNLKVLQRALQDTP 300
QY 316 VDEGGWMLPQETLESGSRLEQ 337
Db 301 VDEGGWMLPQETLESGSRLEQ 322

RESULT 3

US-09-995-225-20
Sequence 20, Application US/09995225
Publication No. US20030139588A9
GENERAL INFORMATION:
APPLICANT: Chen, Ruoping
APPLICANT: Chu, Zhi Liang
APPLICANT: Dang, Huong T.
APPLICANT: Lowitz, Kevin P.
APPLICANT: Pride, Cameron
TITLE OF INVENTION: Endogenous And No. US20030139588A9-Endogenous Versions of Human G
FILE REFERENCE: AREN-0308
CURRENT APPLICATION NUMBER: US/09/995,225
CURRENT FILING DATE: 2001-11-26
PRIOR APPLICATION NUMBER: 09/170,496
PRIOR FILING DATE: 1998-10-13
PRIOR APPLICATION NUMBER: PCT/US99/23938
PRIOR FILING DATE: 1998-10-13
PRIOR APPLICATION NUMBER: 60/253,404
PRIOR FILING DATE: 2000-11-27
PRIOR APPLICATION NUMBER: 60/255,366
PRIOR FILING DATE: 2000-12-12
PRIOR APPLICATION NUMBER: 60/270,286
PRIOR FILING DATE: 2001-02-20
PRIOR APPLICATION NUMBER: 60/282,365
PRIOR FILING DATE: 2001-04-06
PRIOR APPLICATION NUMBER: 60/270,266
PRIOR FILING DATE: 2001-02-20
PRIOR APPLICATION NUMBER: 60/282,032
PRIOR FILING DATE: 2001-04-06
PRIOR APPLICATION NUMBER: 60/282,358
PRIOR FILING DATE: 2001-04-06
PRIOR APPLICATION NUMBER: 60/282,356
PRIOR FILING DATE: 2001-04-06
PRIOR APPLICATION NUMBER: 60/290,917
PRIOR FILING DATE: 2001-05-14
PRIOR APPLICATION NUMBER: 60/309,208
PRIOR FILING DATE: 2001-07-31
NUMBER OF SEQ ID NOS: 67
SOFTWARE: PatentIn version 3.1
SEQ ID NO 20
LENGTH: 322
TYPE: PRT
ORGANISM: Artificial Sequence
FEATURE:
OTHER INFORMATION: No. US20030139588A9 Sequence
US-09-995-225-20

Query Match 95.7%; Score 1688; DB 10; Length 322;
Best Local Similarity 100.0%; Pred. No. 4.6e-145;
Matches 322; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 16 MDSTIPVLGTELPINGRETPCYKOTLSFTGLTCTIVSLVATLGNVAVMLLGCRRRNA 75
DB 1 MDSTIPVLGTELPINGRETPCYKOTLSFTGLTCTIVSLVATLGNVAVMLLGCRRRNA 60
QY 76 VSIYIINLVAADEFLSGHIICSPRLINIRHPISKILSPVMTFPYFGLSMISAISTER 135
DB 61 VSIYIINLVAADEFLSGHIICSPRLINIRHPISKILSPVMTFPYFGLSMISAISTER 120
QY 136 CUSILMPPIWYHCRPRYLSSVMCVLLMALSLRSILEMFCDFLFGADSVWCETSDFTT 195
DB 121 CUSILMPPIWYHCRPRYLSSVMCVLLMALSLRSILEMFCDFLFGADSVWCETSDFTT 180
QY 196 IAWLVELCVLCCSSVLVLRILCGSRKMPRLRYVTIILLTVLVLGCLPGIOWALFS 255
DB 181 IAWLVELCVLCCSSVLVLRILCGSRKMPRLRYVTIILLTVLVLGCLPGIOWALFS 240
QY 256 RIHLDMKVLFCVHVLVSIPLSALNSSANPIIYFVGSFRORONRMLKVLQALODTPE 315
DB 241 RIHLDMKVLFCVHVLVSIPLSALNSSANPIIYFVGSFRORONRMLKVLQALODTPE 300
QY 316 VDEGGMLPOETLELSGSRLEQ 337
DB 301 VDEGGMLPOETLELSGSRLEQ 322

RESULT 4

US-10-183-116-31
; Sequence 31, Application US/10183116
; Publication No. US2003092035A1
; GENERAL INFORMATION:
; APPLICANT: Anderson, David J.
; APPLICANT: Dong, Xinzhang
; APPLICANT: Zylka, Mark
; APPLICANT: Simon, Melvin
; APPLICANT: Han, Sang-Kyun
; TITLE OF INVENTION: PAIN SIGNALING MOLECULES
; FILE REFERENCE: CALTE.4CICPI
; CURRENT APPLICATION NUMBER: US/10/183,116
; CURRENT FILING DATE: 2002-06-26
; PRIOR APPLICATION NUMBER: US 60/222,344
; PRIOR FILING DATE: 2000-08-01
; PRIOR APPLICATION NUMBER: US 60/202,027
; PRIOR FILING DATE: 2000-05-04
; PRIOR APPLICATION NUMBER: US 09/704,707
; PRIOR FILING DATE: 2000-11-03
; PRIOR APPLICATION NUMBER: US 60/285,493
; PRIOR FILING DATE: 2001-04-19
; PRIOR APPLICATION NUMBER: US 09/849,869
; PRIOR FILING DATE: 2001-05-04
; NUMBER OF SEQ ID NOS: 109
; SOFTWARE: PatSeq for Windows Version 4.0
; SEQ ID NO 31
; LENGTH: 322
; TYPE: PRT
; ORGANISM: Homo sapiens
US-10-183-116-31

Query Match 95.7%; Score 1688; DB 14; Length 322;
Best Local Similarity 100.0%; Pred. No. 4,6e-145;
Matches 322; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 16 MDSTIPVLGTELPINGRETPCYKOTLSFTGLTCTIVSLVATLGNVAVMLLGCRRRNA 75
DB 1 MDSTIPVLGTELPINGRETPCYKOTLSFTGLTCTIVSLVATLGNVAVMLLGCRRRNA 60
QY 76 VSIYIINLVAADEFLSGHIICSPRLINIRHPISKILSPVMTFPYFGLSMISAISTER 135
DB 61 VSIYIINLVAADEFLSGHIICSPRLINIRHPISKILSPVMTFPYFGLSMISAISTER 120
QY 136 CUSILMPPIWYHCRPRYLSSVMCVLLMALSLRSILEMFCDFLFGADSVWCETSDFTT 195
DB 121 CUSILMPPIWYHCRPRYLSSVMCVLLMALSLRSILEMFCDFLFGADSVWCETSDFTT 180

QY 196 IAWLVELCVLCCSSVLVLRILCGSRKMPRLRYVTIILLTVLVLGCLPGIOWALFS 255
DB 181 IAWLVELCVLCCSSVLVLRILCGSRKMPRLRYVTIILLTVLVLGCLPGIOWALFS 240
QY 256 RIHLDMKVLFCVHVLVSIPLSALNSSANPIIYFVGSFRORONRMLKVLQALODTPE 315
DB 241 RIHLDMKVLFCVHVLVSIPLSALNSSANPIIYFVGSFRORONRMLKVLQALODTPE 300
QY 316 VDEGGMLPOETLELSGSRLEQ 337
DB 301 VDEGGMLPOETLELSGSRLEQ 322

RESULT 5

US-10-225-567A-674
; Sequence 674, Application US/10225567A
; Publication No. US20030113798A1
; GENERAL INFORMATION:
; APPLICANT: LifeSpan Biosciences
; APPLICANT: Brown, Joseph P.
; APPLICANT: Burner, Glenn C.
; APPLICANT: Roush, Christine L.
; TITLE OF INVENTION: ANTIGENIC PEPTIDES AND ANTIBODIES FOR G PROTEIN-COUPLED RECEPTORS
; FILE REFERENCE: 1920-4-4
; CURRENT APPLICATION NUMBER: US/10/225,567A
; CURRENT FILING DATE: 2001-12-19
; PRIOR APPLICATION NUMBER: 60/257,144
; PRIOR FILING DATE: 2000-12-19
; NUMBER OF SEQ ID NOS: 2292
; SOFTWARE: PatentIn version 3.1
; SEQ ID NO 674
; LENGTH: 322
; TYPE: PRT
; ORGANISM: Homo sapiens
US-10-225-567A-674

Query Match 95.7%; Score 1688; DB 14; Length 322;
Best Local Similarity 100.0%; Pred. No. 4,6e-145;
Matches 322; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 16 MDSTIPVLGTELPINGRETPCYKOTLSFTGLTCTIVSLVATLGNVAVMLLGCRRRNA 75
DB 1 MDSTIPVLGTELPINGRETPCYKOTLSFTGLTCTIVSLVATLGNVAVMLLGCRRRNA 60
QY 76 VSIYIINLVAADEFLSGHIICSPRLINIRHPISKILSPVMTFPYFGLSMISAISTER 135
DB 61 VSIYIINLVAADEFLSGHIICSPRLINIRHPISKILSPVMTFPYFGLSMISAISTER 120
QY 136 CUSILMPPIWYHCRPRYLSSVMCVLLMALSLRSILEMFCDFLFGADSVWCETSDFTT 195
DB 121 CUSILMPPIWYHCRPRYLSSVMCVLLMALSLRSILEMFCDFLFGADSVWCETSDFTT 180
QY 196 IAWLVELCVLCCSSVLVLRILCGSRKMPRLRYVTIILLTVLVLGCLPGIOWALFS 255
DB 181 IAWLVELCVLCCSSVLVLRILCGSRKMPRLRYVTIILLTVLVLGCLPGIOWALFS 240
QY 256 RIHLDMKVLFCVHVLVSIPLSALNSSANPIIYFVGSFRORONRMLKVLQALODTPE 315
DB 241 RIHLDMKVLFCVHVLVSIPLSALNSSANPIIYFVGSFRORONRMLKVLQALODTPE 300
QY 316 VDEGGMLPOETLELSGSRLEQ 337
DB 301 VDEGGMLPOETLELSGSRLEQ 322

RESULT 6

US-10-072-012-529
; Sequence 529, Application US/10072012
; Publication No. US20040033493A1
; GENERAL INFORMATION:
; APPLICANT: Tchernev, Velizar
; APPLICANT: Spytek, Kimberly
; APPLICANT: Zerhusen, Bryan

APPLICANT: Patturajan, Meera
APPLICANT: Shimkets, Richard
APPLICANT: Li, Li
APPLICANT: Gangolli, Esha
APPLICANT: Padigaru, Muralidhara
APPLICANT: Anderson, David W.
APPLICANT: Rastelli, Luca
APPLICANT: Miller, Charles E.
APPLICANT: Gerlach, Valerie
APPLICANT: Taupier Jr, Raymond J.
APPLICANT: Gusev, Vladimir Y.
APPLICANT: Colman, Steven D.
APPLICANT: Wolenc, Adam R.
APPLICANT: Pena, Carol E. A
APPLICANT: Furtak, Katarzyna
APPLICANT: Grose, William M.
APPLICANT: Alsebrook II, John P.
APPLICANT: Lepley, Denise M.
APPLICANT: Rieger, Daniel K.
APPLICANT: Burgess, Catherine E.
TITLE OF INVENTION: Proteins and Nucleic Acids Encoding Same
FILE REFERENCE: 21402-258
CURRENT FILING DATE: 2002-01-31
PRIOR APPLICATION NUMBER: 60/265,102
PRIOR FILING DATE: 2001-01-30
PRIOR APPLICATION NUMBER: 60/265,514
PRIOR FILING DATE: 2001-01-31
PRIOR APPLICATION NUMBER: 60/265,517
PRIOR FILING DATE: 2001-01-31
PRIOR APPLICATION NUMBER: 60/265,412
PRIOR FILING DATE: 2001-01-31
PRIOR APPLICATION NUMBER: 60/265,395
PRIOR FILING DATE: 2001-01-31
PRIOR APPLICATION NUMBER: 60/266,406
PRIOR FILING DATE: 2001-02-02
PRIOR APPLICATION NUMBER: 60/266,767
PRIOR FILING DATE: 2001-02-05
PRIOR APPLICATION NUMBER: 60/267,057
PRIOR FILING DATE: 2001-02-07
PRIOR APPLICATION NUMBER: 60/266,975
PRIOR FILING DATE: 2001-02-07
PRIOR APPLICATION NUMBER: 60/267,459
PRIOR FILING DATE: 2001-02-08
Remaining Prior Application data removed - See file Wrapper or PALM.
NUMBER OF SEQ ID NOS: 1391
SOFTWARE: PatentIn Ver. 2.1
SEQ ID NO 529
LENGTH: 322
TYPE: PRT
ORGANISM: Homo sapiens
US-10-072-012-529

Query Match 95.7%; Score 1688; DB 15; Length 322;
Best Local Similarity 100.0%; Pred. No. 4.6e-145;

Matches 322; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

16 MDSTIPVAGTELTPTINGREPTCYKQTSFTGLCTIVSLVLTGNAVLMILGCMRANA 75
1 MDSTIPVAGTELTPTINGREPTCYKQTSFTGLCTIVSLVLTGNAVLMILGCMRANA 60
61 VSIYILNVADFLFLSGHII CSPRLINIRHPISKILSPVTFPIGLSMLAISTER 120
76 VSIYILNVADFLFLSGHII CSPRLINIRHPISKILSPVTFPIGLSMLAISTER 135
136 CLSTIPWTHCRPRYSVWCVLMLSLRSLTLEMFCDFLFGADSVWCERTSDPT 195
121 CLSTIPWTHCRPRYSVWCVLMLSLRSLTLEMFCDFLFGADSVWCERTSDPT 180
196 IAWIVPLCVVCGSSLVLLVRLGSRKMPRLTKVTTILTVLVLGSLPGIOMALFS 255
181 IAWIVPLCVVCGSSLVLLVRLGSRKMPRLTKVTTILTVLVLGSLPGIOMALFS 240

256 RIHDMKYLFCCHVHVSIFLSALNSSANPITTFPVGSFRORORONLKVLRALODTPE 315
241 RIHDMKYLFCCHVHVSIFLSALNSSANPITTFPVGSFRORORONLKVLRALODTPE 300
316 VDEGGWLPORTELSSGRLEQ 337
301 VDEGGWLPORTELSSGRLEQ 322

RESULT 7
US-10-072-012-534

Sequence 534; Application US/10072012
Publication No. US2004003493A1

GENERAL INFORMATION:

APPLICANT: Tchernev, Velizar

APPLICANT: Spytek, Kimberly

APPLICANT: Zernusen, Bryan

APPLICANT: Patturajan, Meera

APPLICANT: Shimkets, Richard

APPLICANT: Li, Li

APPLICANT: Gangolli, Esha

APPLICANT: Padigaru, Muralidhara

APPLICANT: Anderson, David W.

APPLICANT: Rastelli, Luca

APPLICANT: Miller, Charles E.

APPLICANT: Gerlach, Valerie

APPLICANT: Taupier Jr, Raymond J.

APPLICANT: Gusev, Vladimir Y.

APPLICANT: Colman, Steven D.

APPLICANT: Wolenc, Adam R.

APPLICANT: Pena, Carol E. A

APPLICANT: Furtak, Katarzyna

APPLICANT: Grose, William M.

APPLICANT: Alsebrook II, John P.

APPLICANT: Lepley, Denise M.

APPLICANT: Rieger, Daniel K.

APPLICANT: Burgess, Catherine E.

TITLE OF INVENTION: Proteins and Nucleic Acids Encoding Same

FILE REFERENCE: 21402-258

CURRENT FILING DATE: 2002-01-31

PRIOR APPLICATION NUMBER: 60/265,102

PRIOR FILING DATE: 2001-01-30

PRIOR APPLICATION NUMBER: 60/265,514

PRIOR FILING DATE: 2001-01-31

PRIOR APPLICATION NUMBER: 60/265,517

PRIOR FILING DATE: 2001-01-31

PRIOR APPLICATION NUMBER: 60/265,412

PRIOR FILING DATE: 2001-01-31

PRIOR APPLICATION NUMBER: 60/265,395

PRIOR FILING DATE: 2001-01-31

PRIOR APPLICATION NUMBER: 60/266,406

PRIOR FILING DATE: 2001-02-02

PRIOR APPLICATION NUMBER: 60/266,767

PRIOR FILING DATE: 2001-02-05

PRIOR APPLICATION NUMBER: 60/267,057

PRIOR FILING DATE: 2001-02-07

PRIOR APPLICATION NUMBER: 60/266,975

PRIOR FILING DATE: 2001-02-07

PRIOR APPLICATION NUMBER: 60/267,459

PRIOR FILING DATE: 2001-02-08

Remaining Prior Application data removed - See file Wrapper or PALM.

NUMBER OF SEQ ID NOS: 1391

SOFTWARE: PatentIn Ver. 2.1

SEQ ID NO 534

LENGTH: 322

TYPE: PRT

ORGANISM: Homo sapiens

US-10-072-012-534

Query Match 95.7%; Score 1688; DB 15; Length 322;
Best Local Similarity 100.0%; Pred. No. 4.6e-145;

Matches 322; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 16 MDSTIPVLGTELTPIINGREBETPCYKOTLSTFTGLTCTIVSLVALTGNNAVLMGLGCRMRNA 75
DB 1 MDSTIPVLGTELTPIINGREBETPCYKOTLSTFTGLTCTIVSLVALTGNNAVLMGLGCRMRNA 60
QY 76 VSIYIINLVAADPLFLSGHIICSPRLINIRHPISKILSPVMTFPYFIFGLSMISAISTER 135
DB 61 VSIYIINLVAADPLFLSGHIICSPRLINIRHPISKILSPVMTFPYFIFGLSMISAISTER 120
QY 136 CSTIIMPWYHCHRRPRLSSVWCVLMLALSILRSILEMFCDFLFGSADSVWCETSDPFT 195
DB 121 CSTIIMPWYHCHRRPRLSSVWCVLMLALSILRSILEMFCDFLFGSADSVWCETSDPFT 180
QY 196 IAWLVEFLCVLCCSSVLVLRILCGSRKMPRLRYTYTILLTVLVFLCGLPFGIQWALFS 255
DB 181 IAWLVEFLCVLCCSSVLVLRILCGSRKMPRLRYTYTILLTVLVFLCGLPFGIQWALFS 240
QY 256 RIHLDMKVLFCVHVLVSIPLSALNSSANPIIYFVGSFRORONRQMLKVLORALODTPE 315
DB 241 RIHLDMKVLFCVHVLVSIPLSALNSSANPIIYFVGSFRORONRQMLKVLORALODTPE 300
QY 316 VDEGGGMLPOETTELSSGRLAQ 337
DB 301 VDEGGGMLPOETTELSSGRLAQ 322

RESULT 8

US-10-401-397A-2
; Sequence 2, Application US/10401397A
; Publication No. US20030221001A1
; GENERAL INFORMATION:
; APPLICANT: Peri, Krishna G.
; APPLICANT: Mofect, Serge
; APPLICANT: Abiram, Daniel
; TITLE OF INVENTION: METHODS AND COMPOUNDS FOR PREVENTION AND TREATMENT OF ELEVATED
; TITLE OF INVENTION: INTRAOCULAR PRESSURE AND RELATED CONDITIONS
; FILE REFERENCE: 4518/1M674US1
; CURRENT APPLICATION NUMBER: US/10/401,397A
; PRIOR FILING DATE: 2003-03-27
; PRIOR APPLICATION NUMBER: US 60/367,513
; PRIOR FILING DATE: 2002-03-27
; NUMBER OF SEQ ID NOS: 8
; SOFTWARE: PatentIn version 3.1
; SEQ ID NO 2
; LENGTH: 322
; TYPE: PRT
; ORGANISM: Homo sapiens
US-10-401-397A-2

Query Match 95.5%; Score 1683; DB 14; Length 322;

Best Local Similarity 99.7%; Pred. No. 1.3e-144; Matches 321; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

QY 16 MDSTIPVLGTELTPIINGREBETPCYKOTLSTFTGLTCTIVSLVALTGNNAVLMGLGCRMRNA 75
DB 1 MDSTIPVLGTELTPIINGREBETPCYKOTLSTFTGLTCTIVSLVALTGNNAVLMGLGCRMRNA 60
QY 76 VSIYIINLVAADPLFLSGHIICSPRLINIRHPISKILSPVMTFPYFIFGLSMISAISTER 135
DB 61 VSIYIINLVAADPLFLSGHIICSPRLINIRHPISKILSPVMTFPYFIFGLSMISAISTER 120
QY 136 CSTIIMPWYHCHRRPRLSSVWCVLMLALSILRSILEMFCDFLFGSADSVWCETSDPFT 195
DB 121 CSTIIMPWYHCHRRPRLSSVWCVLMLALSILRSILEMFCDFLFGSADSVWCETSDPFT 180
QY 196 IAWLVEFLCVLCCSSVLVLRILCGSRKMPRLRYTYTILLTVLVFLCGLPFGIQWALFS 255
DB 181 IAWLVEFLCVLCCSSVLVLRILCGSRKMPRLRYTYTILLTVLVFLCGLPFGIQWALFS 240
QY 256 RIHLDMKVLFCVHVLVSIPLSALNSSANPIIYFVGSFRORONRQMLKVLORALODTPE 315
DB 241 RIHLDMKVLFCVHVLVSIPLSALNSSANPIIYFVGSFRORONRQMLKVLORALODTPE 300

QY 316 VDEGGGMLPOETTELSSGRLAQ 337
DB 301 VDEGGGMLPOETTELSSGRLAQ 322

RESULT 9

US-10-292-798-1274
; Sequence 1274, Application US/10292798
; Publication No. US2003023583A1
; GENERAL INFORMATION:
; APPLICANT: SUWA, MAKIKO
; APPLICANT: ASAI, KIYOSHI
; APPLICANT: AKIYAMA, YUTAKA
; APPLICANT: AUBURATANI, HIROYUKI
; TITLE OF INVENTION: GUANOSINE TRIPHOSPHATE-BINDING PROTEIN COUPLED RECEPTORS
; FILE REFERENCE: 08435/166
; CURRENT APPLICATION NUMBER: US/10/292,798
; PRIOR FILING DATE: 2002-11-13
; PRIOR APPLICATION NUMBER: 10/017,161
; PRIOR FILING DATE: 2001-12-18
; PRIOR APPLICATION NUMBER: JP 2001-246789
; PRIOR FILING DATE: 2001-06-18
; NUMBER OF SEQ ID NOS: 2070
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 1274
; LENGTH: 322
; TYPE: PRT
; ORGANISM: Homo sapiens
US-10-292-798-1274

Query Match 95.5%; Score 1683; DB 14; Length 322;

Best Local Similarity 99.7%; Pred. No. 1.3e-144; Matches 321; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

QY 16 MDSTIPVLGTELTPIINGREBETPCYKOTLSTFTGLTCTIVSLVALTGNNAVLMGLGCRMRNA 75
DB 1 MDSTIPVLGTELTPIINGREBETPCYKOTLSTFTGLTCTIVSLVALTGNNAVLMGLGCRMRNA 60
QY 76 VSIYIINLVAADPLFLSGHIICSPRLINIRHPISKILSPVMTFPYFIFGLSMISAISTER 135
DB 61 VSIYIINLVAADPLFLSGHIICSPRLINIRHPISKILSPVMTFPYFIFGLSMISAISTER 120
QY 136 CSTIIMPWYHCHRRPRLSSVWCVLMLALSILRSILEMFCDFLFGSADSVWCETSDPFT 195
DB 121 CSTIIMPWYHCHRRPRLSSVWCVLMLALSILRSILEMFCDFLFGSADSVWCETSDPFT 180
QY 196 IAWLVEFLCVLCCSSVLVLRILCGSRKMPRLRYTYTILLTVLVFLCGLPFGIQWALFS 255
DB 181 IAWLVEFLCVLCCSSVLVLRILCGSRKMPRLRYTYTILLTVLVFLCGLPFGIQWALFS 240
QY 256 RIHLDMKVLFCVHVLVSIPLSALNSSANPIIYFVGSFRORONRQMLKVLORALODTPE 315
DB 241 RIHLDMKVLFCVHVLVSIPLSALNSSANPIIYFVGSFRORONRQMLKVLORALODTPE 300
QY 316 VDEGGGMLPOETTELSSGRLAQ 337
DB 301 VDEGGGMLPOETTELSSGRLAQ 322

RESULT 10

US-10-072-012-530
; Sequence 530, Application US/10072012
; Publication No. US20040033493A1
; GENERAL INFORMATION:
; APPLICANT: Tchernev, Velizar
; APPLICANT: Spytek, Kimberly
; APPLICANT: Zerkusen, Bryan
; APPLICANT: Patirajan, Meera
; APPLICANT: Shinkets, Richard
; APPLICANT: Li, Li
; APPLICANT: Gangolli, Saba
; APPLICANT: Padigaru, Muralidhara
; APPLICANT: Anderson, David W.

```

/ APPLICANT: Raetelli, Luca
/ APPLICANT: Miller, Charles E.
/ APPLICANT: Gerlach, Valerie
/ APPLICANT: Taupier Jr, Raymond J.
/ APPLICANT: Gusev, Vladimir Y.
/ APPLICANT: Coleman, Steven D.
/ APPLICANT: Wolenc, Adam R.
/ APPLICANT: Pena, Carol E. A
/ APPLICANT: Furtak, Katarzyna
/ APPLICANT: Grose, William M.
/ APPLICANT: Alsbrook II, John P.
/ APPLICANT: Lepley, Denise M.
/ APPLICANT: Rieger, Daniel K.
/ APPLICANT: Burgess, Catherine E.
/ TITLE OF INVENTION: Proteins and Nucleic Acids Encoding Same
/ FILE REFERENCE: 21402-258
/ CURRENT APPLICATION NUMBER: US/10/072,012
/ CURRENT FILING DATE: 2002-01-31
/ PRIOR APPLICATION NUMBER: 60/265,102
/ PRIOR FILING DATE: 2001-01-30
/ PRIOR APPLICATION NUMBER: 60/265,514
/ PRIOR FILING DATE: 2001-01-31
/ PRIOR APPLICATION NUMBER: 60/265,517
/ PRIOR FILING DATE: 2001-01-31
/ PRIOR APPLICATION NUMBER: 60/265,412
/ PRIOR FILING DATE: 2001-01-31
/ PRIOR APPLICATION NUMBER: 60/265,395
/ PRIOR FILING DATE: 2001-01-31
/ PRIOR APPLICATION NUMBER: 60/266,406
/ PRIOR FILING DATE: 2001-02-02
/ PRIOR APPLICATION NUMBER: 60/266,767
/ PRIOR FILING DATE: 2001-02-05
/ PRIOR APPLICATION NUMBER: 60/267,057
/ PRIOR FILING DATE: 2001-02-07
/ PRIOR APPLICATION NUMBER: 60/266,975
/ PRIOR FILING DATE: 2001-02-07
/ PRIOR APPLICATION NUMBER: 60/267,459
/ PRIOR FILING DATE: 2001-02-08
/ Remaining Prior Application data removed - See File Wrapper or PALM.
/ NUMBER OF SEQ ID NOS: 1391
/ SOFTWARE: PatentIn Ver. 2.1
/ SEQ ID NO 530
/ LENGTH: 322
/ TYPE: PRT
/ ORGANISM: Homo sapiens
/ US-10-072-012-530

Query Match          95.5%; Score 1683; DB 15; Length 322;
Best Local Similarity 99.7%; Pred. No. 1.3e-144;
Matches 321; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

```

```

Db          301 VDEGGWLPORTELESGSRLEQ 322

RESULT 11
US-10-072-012-535
/ Sequence 535, Application US/10072012
/ Publication No. US2004003493A1
/ GENERAL INFORMATION:
/ APPLICANT: Tchernev, Velizar
/ APPLICANT: Spytek, Kimberly
/ APPLICANT: Zerkusen, Bryan
/ APPLICANT: Patturajan, Meera
/ APPLICANT: Shimkets, Richard
/ APPLICANT: Li, Li
/ APPLICANT: Gangoli, Esha
/ APPLICANT: Padigaru, Muralidhara
/ APPLICANT: Anderson, David W.
/ APPLICANT: Raetelli, Luca
/ APPLICANT: Miller, Charles E.
/ APPLICANT: Gerlach, Valerie
/ APPLICANT: Taupier Jr, Raymond J.
/ APPLICANT: Gusev, Vladimir Y.
/ APPLICANT: Coleman, Steven D.
/ APPLICANT: Wolenc, Adam R.
/ APPLICANT: Pena, Carol E. A
/ APPLICANT: Furtak, Katarzyna
/ APPLICANT: Grose, William M.
/ APPLICANT: Alsbrook II, John P.
/ APPLICANT: Lepley, Denise M.
/ APPLICANT: Rieger, Daniel K.
/ APPLICANT: Burgess, Catherine E.
/ TITLE OF INVENTION: Proteins and Nucleic Acids Encoding Same
/ FILE REFERENCE: 21402-258
/ CURRENT APPLICATION NUMBER: US/10/072,012
/ CURRENT FILING DATE: 2002-01-31
/ PRIOR APPLICATION NUMBER: 60/265,102
/ PRIOR FILING DATE: 2001-01-30
/ PRIOR APPLICATION NUMBER: 60/265,514
/ PRIOR FILING DATE: 2001-01-31
/ PRIOR APPLICATION NUMBER: 60/265,517
/ PRIOR FILING DATE: 2001-01-31
/ PRIOR APPLICATION NUMBER: 60/265,412
/ PRIOR FILING DATE: 2001-01-31
/ PRIOR APPLICATION NUMBER: 60/265,395
/ PRIOR FILING DATE: 2001-01-31
/ PRIOR APPLICATION NUMBER: 60/266,406
/ PRIOR FILING DATE: 2001-02-02
/ PRIOR APPLICATION NUMBER: 60/266,767
/ PRIOR FILING DATE: 2001-02-05
/ PRIOR APPLICATION NUMBER: 60/267,057
/ PRIOR FILING DATE: 2001-02-07
/ PRIOR APPLICATION NUMBER: 60/266,975
/ PRIOR FILING DATE: 2001-02-07
/ PRIOR APPLICATION NUMBER: 60/267,459
/ PRIOR FILING DATE: 2001-02-08
/ Remaining Prior Application data removed - See File Wrapper or PALM.
/ NUMBER OF SEQ ID NOS: 1391
/ SOFTWARE: PatentIn Ver. 2.1
/ SEQ ID NO 535
/ LENGTH: 322
/ TYPE: PRT
/ ORGANISM: Homo sapiens
/ US-10-072-012-535

Query Match          95.5%; Score 1683; DB 15; Length 322;
Best Local Similarity 99.7%; Pred. No. 1.3e-144;
Matches 321; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

```

DB 61 VSIIYIINLVAADEFLFSGHIIICSPRLINIRHPIISKILSPVMTFPYFISGLMSAISTER 120
QY 136 CUSTLMPWYHCRPRYLSVWCVLMLALSILSIEMWFCDFLFGSADSVWCETSDPIT 195
DB 121 CUSTLMPWYHCRPRYLSVWCVLMLALSILSIEMWFCDFLFGSADSVWCETSDPIT 180
QY 196 IAWLVLCVLCGSSVLVLRILICGSRKMPLTRLYVTIILTVAVPLLCGLPGIOWALFS 255
DB 181 IAWLVLCVLCGSSVLVLRILICGSRKMPLTRLYVTIILTVAVPLLCGLPGIOWALFS 240
QY 256 RIHLDMKVLFCVHVLVSIPLSALNSSANPIIYFVGSFRORONRQNLKVLQALODTPE 315
DB 241 RIHLDMKVLFCVHVLVSIPLSALNSSANPIIYFVGSFRORONRQNLKVLQALODTPE 300
QY 316 VDEGGGMLPOETIELSGSRLEQ 337
DB 301 VDEGGGMLPOETIELSGSRLEQ 322

RESULT 12

US-10-391-074-2
; Sequence 2, Application US/10391074
; Publication No. US20040038345A1
; GENERAL INFORMATION:
; APPLICANT: Nehls, Michael
; APPLICANT: Matlier, Frank
; TITLE OF INVENTION: No. US20040038345A1 Human Seven-Transmembrane Receptors
; FILE REFERENCE: 7705-0008-00-000
; CURRENT APPLICATION NUMBER: US/10/391,074
; NUMBER OF SEQ ID NOS: 4
; SOFTWARE: FastSeq for Windows Version 3.0
; SEQ ID NO 2
; LENGTH: 322
; TYPE: PRT
; ORGANISM: Homo sapien
US-10-391-074-2

Query Match 94.3%; Score 1663; DB 15; Length 322;

Best Local Similarity 99.4%; Pred. No. 8.5e-143;

Matches 330; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

QY 16 MDSTIPVLGTELPIPINGRETEPCYKOTLSFTGLTCTIVSVALTGNNAVLMILGCRMRNA 75
DB 1 MDSTIPVLGTELPIPINGRETEPCYKOTLSFTGLTCTIVSVALTGNNAVLMILGCRMRNA 60
QY 76 VSIIYIINLVAADEFLFSGHIIICSPRLINIRHPIISKILSPVMTFPYFISGLMSAISTER 135
DB 61 VSIIYIINLVAADEFLFSGHIIIRSPRLINIRHPIISKILSPVMTFPYFISGLMSAISTER 120
QY 136 CUSTLMPWYHCRPRYLSVWCVLMLALSILSIEMWFCDFLFGSADSVWCETSDPIT 195
DB 121 CUSTLMPWYHCRPRYLSVWCVLMLALSILSIEMWFCDFLFGSADSVWCETSDPIT 180
QY 196 IAWLVLCVLCGSSVLVLRILICGSRKMPLTRLYVTIILTVAVPLLCGLPGIOWALFS 255
DB 181 IAWLVLCVLCGSSVLVLRILICGSRKMPLTRLYVTIILTVAVPLLCGLPGIOWALFS 240
QY 256 RIHLDMKVLFCVHVLVSIPLSALNSSANPIIYFVGSFRORONRQNLKVLQALODTPE 315
DB 241 RIHLDMKVLFCVHVLVSIPLSALNSSANPIIYFVGSFRORONRQNLKVLQALODTPE 300
QY 316 VDEGGGMLPOETIELSGSRLEQ 337
DB 301 VDEGGGMLPOETIELSGSRLEQ 322

RESULT 13

US-10-219-834-79
; Sequence 79, Application US/10219834
; Publication No. US20030096751A1
; GENERAL INFORMATION:

; APPLICANT: Bristol-Myers Squibb Company
; TITLE OF INVENTION: G-PROTEIN COUPLED RECEPTOR POLYNUCLEOTIDES AND METHODS OF USE THEREOF
; FILE REFERENCE: D0191 NP
; CURRENT APPLICATION NUMBER: US/10/219,834
; PRIOR FILING DATE: 2002-08-15
; PRIOR APPLICATION NUMBER: US 60/313,658
; PRIOR FILING DATE: 2001-08-20
; PRIOR APPLICATION NUMBER: US 60/340,703
; PRIOR FILING DATE: 2001-10-30
; PRIOR APPLICATION NUMBER: US 60/318,675
; PRIOR FILING DATE: 2001-09-12
; PRIOR APPLICATION NUMBER: US 60/355,596
; PRIOR FILING DATE: 2002-02-06
; PRIOR APPLICATION NUMBER: US 60/333,417
; PRIOR FILING DATE: 2001-11-26
; PRIOR APPLICATION NUMBER: US 60/338,367
; PRIOR FILING DATE: 2001-12-06
; NUMBER OF SEQ ID NOS: 192
; SOFTWARE: PatentIn version 3.1
; SEQ ID NO 79
; LENGTH: 314
; TYPE: PRT
; ORGANISM: Homo sapiens
US-10-219-834-79

Query Match 93.0%; Score 1639; DB 14; Length 314;
Best Local Similarity 97.5%; Pred. No. 1.3e-140;
Matches 314; Conservative 0; Mismatches 0; Indels 8; Gaps 1;

QY 16 MDSTIPVLGTELPIPINGRETEPCYKOTLSFTGLTCTIVSVALTGNNAVLMILGCRMRNA 75
DB 1 MDSTIPVLGTELPIPINGRETEPCYKOTLSFTGLTCTIVSVALTGNNAVLMILGCRMRNA 60
QY 76 VSIIYIINLVAADEFLFSGHIIICSPRLINIRHPIISKILSPVMTFPYFISGLMSAISTER 135
DB 61 VSIIYIINLVAADEFLFSGHIIICSPRLINIRHPIISKILSPVMTFPYFISGLMSAISTER 120
QY 136 CUSTLMPWYHCRPRYLSVWCVLMLALSILSIEMWFCDFLFGSADSVWCETSDPIT 195
DB 121 CUSTLMPWYHCRPRYLSVWCVLMLALSILSIEMWFCDFLFGSADSVWCETSDPIT 180
QY 196 IAWLVLCVLCGSSVLVLRILICGSRKMPLTRLYVTIILTVAVPLLCGLPGIOWALFS 255
DB 181 IAWLVLCVLCGSSVLVLRILICGSRKMPLTRLYVTIILTVAVPLLCGLPGIOWALFS 240
QY 256 RIHLDMKVLFCVHVLVSIPLSALNSSANPIIYFVGSFRORONRQNLKVLQALODTPE 315
DB 241 RIHLDMKVLFCVHVLVSIPLSALNSSANPIIYFVGSFRORONRQNLKVLQALODTPE 300
QY 316 VDEGGGMLPOETIELSGSRLEQ 337
DB 293 VDEGGGMLPOETIELSGSRLEQ 314

RESULT 14

US-10-237-467-10
; Sequence 10, Application US/10237467
; Publication No. US20030186324A1
; GENERAL INFORMATION:
; APPLICANT: Liao, JiaYu
; APPLICANT: Gray, Nathanael S.
; APPLICANT: Caldwell, Jeremy C.
; APPLICANT: Schultz, Peter G.
; APPLICANT: IRM LLC
; TITLE OF INVENTION: Sensory Neuron Receptors
; FILE REFERENCE: 021288-001300US
; CURRENT APPLICATION NUMBER: US/10/237,467
; PRIOR FILING DATE: 2003-01-14
; PRIOR APPLICATION NUMBER: US 60/317,879
; PRIOR FILING DATE: 2001-09-07
; NUMBER OF SEQ ID NOS: 18
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 10

```
; LENGTH: 302
; TYPE: PRT
; ORGANISM: Homo sapiens
; FEATURE:
; OTHER INFORMATION: dorsal root ganglia G-protein coupled receptor (GPCR)
; OTHER INFORMATION: 7 (DRG7) (NT009307)
US-10-237-467-10

Query Match      88.3%; Score 1557; DB 14; Length 302;
Best Local Similarity 93.5%; Pred. No. 3.4e-133;
Matches 301; Conservative 1; Mismatches 0; Indels 20; Gaps 1;

QY 16 MDSTIPVLGTELPINGREETPCYKQTSFTGLTCTIVSLVATGNAVLMILGCMRRNA 75
DB 1 MDSTIPVLGTELPINGREETPCYKQTSFTGLTCTIVSLVATGNAVLMILGCMRRNA 40
QY 76 VSTIILNVLNADPFLSGHIIICSPRLINIRHPSKILSPWTFEYFGLSLAISTER 135
DB 41 VSTIILNVLNADPFLSGHIIICSPRLINIRHPSKILSPWTFEYFGLSLAISTER 100
QY 136 CLSTILMPIMWACRRPRYSVWCVLMLALSLRLSILEMFCDFLFGSGADSVWCETSDFT 195
DB 101 CLSTILMPIMWACRRPRYSVWCVLMLALSLRLSILEMFCDFLFGSGADSVWCETSDFT 160
QY 196 IAWTVFLCVCVLGSSVLVLRILGSSRMQPLRLVTVTLTVLVLGGLPFGIOWALFS 255
DB 161 IAWTVFLCVCVLGSSVLVLRILGSSRMQPLRLVTVTLTVLVLGGLPFGIOWALFS 220
QY 256 RIHDMKVLPCGVHVLVSIFLSALNSSANPIIYFVGSPRORONRQNLKVLORALODPE 315
DB 221 RIHDMKVLPCGVHVLVSIFLSALNSSANPIIYFVGSPRORONRQNLKVLORALODPE 280
QY 316 VDEGGWMLPQETLILSGSRLQ 337
DB 281 VDEGGWMLPQETLILSGSRLQ 302

RESULT 15
US-10-072-012-528
; Sequence 528, Application US/10072012
; Publication No. US20040033493A1
; GENERAL INFORMATION:
; APPLICANT: Tchener, Velizar
; APPLICANT: Spytek, Kimberly
; APPLICANT: Zernhusen, Bryan
; APPLICANT: Paturajan, Meera
; APPLICANT: Shimkets, Richard
; APPLICANT: Li, Li
; APPLICANT: Gangoli, Esha
; APPLICANT: Padigaru, Muralidhara
; APPLICANT: Anderson, David W.
; APPLICANT: Rastelli, Luca
; APPLICANT: Miller, Charles E.
; APPLICANT: Gerlach, Valerie
; APPLICANT: Taupier Jr, Raymond J.
; APPLICANT: Gusev, Vladimir Y.
; APPLICANT: Coleman, Steven D.
; APPLICANT: Wolenc, Adam R.
; APPLICANT: Pena, Carol E. A
; APPLICANT: Furtak, Katarzyna
; APPLICANT: Grose, William M.
; APPLICANT: Alebrock II, John P.
; APPLICANT: Lepley, Denise M.
; APPLICANT: Rieger, Daniel K.
; APPLICANT: Burgess, Catherine E.
; TITLE OR INVENTION: Proteins and Nucleic Acids Encoding Same
; FILE REFERENCE: 21402-258
; CURRENT APPLICATION NUMBER: US/10/072,012
; CURRENT FILING DATE: 2002-01-31
; PRIOR APPLICATION NUMBER: 60/265,102
; PRIOR FILING DATE: 2001-01-30
; PRIOR APPLICATION NUMBER: 60/265,514
; PRIOR FILING DATE: 2001-01-31
```

```
; PRIOR APPLICATION NUMBER: 60/265,517
; PRIOR FILING DATE: 2001-01-31
; PRIOR APPLICATION NUMBER: 60/265,412
; PRIOR FILING DATE: 2001-01-31
; PRIOR APPLICATION NUMBER: 60/265,395
; PRIOR FILING DATE: 2001-01-31
; PRIOR APPLICATION NUMBER: 60/266,406
; PRIOR FILING DATE: 2001-02-02
; PRIOR APPLICATION NUMBER: 60/266,767
; PRIOR FILING DATE: 2001-02-05
; PRIOR APPLICATION NUMBER: 60/267,057
; PRIOR FILING DATE: 2001-02-07
; PRIOR APPLICATION NUMBER: 60/266,975
; PRIOR FILING DATE: 2001-02-07
; PRIOR APPLICATION NUMBER: 60/267,459
; PRIOR FILING DATE: 2001-02-08
; Remaining Prior Application data removed - See File Wrapper or PALM.
; NUMBER OF SEQ ID NOS: 1391
; SOFTWARE: Patent Ver. 2.1
; SEQ ID NO 528
; LENGTH: 1589
; TYPE: PRT
; ORGANISM: Homo sapiens
US-10-072-012-528

Query Match      80.3%; Score 1415.5; DB 15; Length 1589;
Best Local Similarity 79.7%; Pred. No. 1.5e-119;
Matches 279; Conservative 16; Mismatches 38; Indels 17; Gaps 1;

QY 5 SSW-----VIRLGLSMDSTIPVLGTELPINGREETPCYKQTSFTG 47
DB 633 SSWHIKEEHAAGHGVFSRKVTRLGLSMDPTLTSTLTPINGREETPCYKQTSFTV 692
QY 48 LTCIVSLVATGNAVLMILGCMRRNAVSIIYIINLVADPFLSGHIIICSPRLINIRH 107
DB 693 LTCIVSLVATGNAVLMILGCMRRNAFSIIYIINLVADPFLSGHIIICSPRLINIRH 752
QY 108 PISKILSPWTFEYFGLSLAISTERCLSIIMPIMWACRRPRYSVWCVLMLALSLI 167
DB 753 TISKILYFVMMFSYFAGISFLSAVSTERCLSVMPIMWACRRPRYSVWCVLMLALSLI 812
QY 168 RSTLEMMFCDFLFGSGADSVWCETSDFTIAMVFLCVCVLGSSVLVLRILGSSRMQPL 227
DB 813 RSTLEMMFCDFLFGSGADSVWCETSDFTIAMVFLCVCVLGSSVLVLRILGSSRMQPL 872
QY 228 RLVTYIILTVLVLGGLPFGIOWALFSRIHDMKVLPCGVHVLVSIFLSALNSSANPIIY 287
DB 873 RLVTYIILTVLVLGGLPFGIOWALFSRIHDMKVLPCGVHVLVSIFLSALNSSANPIIY 932
QY 288 FVGSPRORONRQNLKVLORALODPEVDEGGWMLPQETLILSGSRLQ 337
DB 933 FVGSPRORONRQNLKVLORALODPEVDEGGWMLPQETLILSGSRLQ 982

Search completed: October 27, 2004, 08:45:33
Job time : 133 secs
```

GenCore version 5.1.6
Copyright (c) 1993 - 2004 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: October 27, 2004, 08:23:27 ; Search time 157 Seconds
(without alignments)
770.011 Million cell updates/sec

Title: US-09-867-570-2

Perfect score: 1763
Sequence: 1 MESKSWVRLGFLSMDSTL.....EGGGMPLQRTLRLSGRLQ 337

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 2002273 seqs, 358729299 residues

Total number of hits satisfying chosen parameters: 2002273

Minimum DB seq length: 0
Maximum DB seq length: 200000000

Post-processing: Minimum Match 0%

Maximum Match 100%
Listing first 45 summaries

Database : A_Geneseq_23Sep04:*

- 1: geneseqp1980s:*
- 2: geneseqp1990s:*
- 3: geneseqp2000s:*
- 4: geneseqp2001s:*
- 5: geneseqp2002s:*
- 6: geneseqp2003as:*
- 7: geneseqp2003bs:*
- 8: geneseqp2004s:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	1763	100.0	337	5	AAU97598 Human G-P
2	1688	95.7	332	3	AA90761 Human G P
3	1688	95.7	322	5	ABJ04077 Human G P
4	1688	95.7	322	5	AAE21296 Human Mrx
5	1688	95.7	322	5	AD116993 Human NOV
6	1688	95.7	322	5	AD116998 Human NOV
7	1688	95.7	322	6	ABP81750 Human G P
8	1688	95.7	322	8	ADH08535 MrGX1_3/
9	1688	95.7	322	8	AD044602 Human HIT
10	1688	95.7	322	8	ADO29705 Human GPC
11	1688	95.7	560	7	ADF70481 Orphan re
12	1683	95.5	322	3	AA90762 Human G P
13	1683	95.5	322	5	AD116994 Human NOV
14	1683	95.5	322	5	AD116999 Human NOV
15	1683	95.5	322	7	AD116999 Human NOV
16	1683	95.5	322	8	AD116999 Human NOV
17	1683	95.5	322	8	AD116999 Human NOV
18	1683	95.5	322	8	AD116999 Human NOV
19	1683	95.5	322	8	AD116999 Human NOV
20	1683	95.5	322	8	AD116999 Human NOV
21	1683	95.5	322	8	AD116999 Human NOV
22	1683	95.5	322	8	AD116999 Human NOV
23	1683	95.5	322	8	AD116999 Human NOV
24	1683	95.5	322	8	AD116999 Human NOV
25	1683	95.5	322	8	AD116999 Human NOV

26	1391	78.9	322	2	AA930162 Human dor
27	1387	78.7	322	3	AA97664 Human G P
28	1387	78.7	322	3	AD117728 Human TGR
29	1383	78.4	322	4	AA64294 Human GPC
30	1383	78.4	322	4	AA612794 Human G P
31	1383	78.4	322	4	AA612794 Human G P
32	1383	78.4	322	5	AA617074 Human G-P
33	1383	78.4	322	5	ABP95617 Human GPC
34	1383	78.4	322	5	AA61288 Human Mrx
35	1383	78.4	322	5	AD116991 Human NOV
36	1383	78.4	322	5	AD116991 Human NOV
37	1383	78.4	322	5	AD116997 Human NOV
38	1383	78.4	322	6	ABP96696 Human G P
39	1383	78.4	322	6	ABP95266 Human GPC
40	1383	78.4	322	7	ABW00803 Human GPC
41	1383	78.4	322	7	AD196466 Human G P
42	1383	78.4	322	8	ADH08520 MrGX1_3/
43	1383	78.4	322	8	ADN42290 Human nov
44	1383	78.4	322	8	ADN42290 Human nov
45	1383	78.4	322	8	ADO29703 Human GPC

ALIGNMENTS

RESULT 1	AAU97598	standard; protein; 337 AA.
ID	AAU97598	
XX	AAU97598;	
AC	12-AUG-2002 (first entry)	
XX		
DT		
XX		
DE	Human G-protein coupled receptor (GPCR).	
XX		
KM	Human; G-protein coupled; receptor; GPCR; human protease;	
KW	human therapeutic protein; query sequence; search; chromosome 3;	
KM	sequence database; non-human transgenic animal; gene therapy.	
XX		
OS	Homo sapiens.	
XX		
PH	Key	Location/Qualifiers
FT	Domain	/label= Helix_1
FT	Domain	/label= Helix_2
FT	Domain	/label= Helix_3
FT	Modified-site	/note= "Protein kinase C (PKC) phosphorylation site"
FT	Domain	/label= Helix_4
FT	Modified-site	/note= "Casein kinase II (CK2) phosphorylation site"
FT	Modified-site	/note= "Casein kinase II (CK2) phosphorylation site"
FT	Domain	/label= Helix_5
FT	Modified-site	/note= "Protein kinase C (PKC) phosphorylation site"
FT	Domain	/label= Helix_6
FT	Modified-site	/note= "Myristoylation site"
FT	Modified-site	/note= "Myristoylation site"
FT	Domain	/label= Helix_7
FT	Modified-site	/note= "Asn glycosylation site"
FT	Modified-site	/note= "Protein kinase C (PKC) phosphorylation site"
FT	Modified-site	/note= "Protein kinase C (PKC) phosphorylation site"

/note= "Casein kinase II (CK2) phosphorylation site"

FT XX WO200234914-A1.
 PN XX 02-MAY-2002.
 PD XX 10-OCT-2001; 2001WO-US031592.
 PF XX 25-OCT-2000; 2000US-00695045.
 PR XX 31-MAY-2001; 2001US-00867570.
 XX (PEKE) PE CORP NY.
 PA Wei M, Zhao Q, Woodage T, Di Francesco V, Beasley EM,
 XX
 PI WPI; 2002-463360/49.
 PN N-PSDB; ABK52822, ABK52823.
 DR
 XX Novel isolated G-protein coupled receptor peptide useful for treating
 PT disorder characterized by absence of, in appropriate or unwanted
 PT expression of the receptor protein, and as immunogens to raise
 PT antibodies.
 XX
 XX Claim 1; Fig 2; 75pp; English.

CC The present invention relates to a new G-protein coupled receptor (GPCR)
 CC peptide. The invention is useful for identifying a modulator of GPCR and
 CC for treating a disease or condition mediated by a human protease. The
 CC invention is also useful as models for the development of human
 CC therapeutics, for identifying therapeutic proteins, as targets for
 CC development of human therapeutic agents, and as query sequence to perform
 CC a search against sequence databases to, for e.g., identify other family
 CC members of related sequences. The vector of the invention is useful for
 CC producing a GPCR protein or peptide, for conducting cell-based assays
 CC involving the GPCR protein or its fragment, for identifying GPCR protein
 CC mutants whose functions are affected, and to produce non-human transgenic
 CC animals. The present amino acid sequence represents the human G-protein
 CC coupled receptor (GPCR) protein of the invention. This sequence is
 CC encoded by the human G-protein coupled receptor (GPCR) gene located on
 CC chromosome 3
 CC
 XX
 SQ Sequence 337 AA;

Query Match 100.0%; Score 1763; DB 5; Length 337;
 Best Local Similarity 100.0%; Pred. No. 1.1e-180;
 Matches 337; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 MESKSSWTRIGELSMDSITPVLGTELPINGREETPCYKQTLSTFGTLCIYSLVALTGN 60
 DB 1 MESKSSWTRIGELSMDSITPVLGTELPINGREETPCYKQTLSTFGTLCIYSLVALTGN 60
 QY 61 AVVLMILGCRMRNNAVSIYIILNVAADFLFSGHIICSPRLINIRHPISKILSPVMP 120
 DB 61 AVVLMILGCRMRNNAVSIYIILNVAADFLFSGHIICSPRLINIRHPISKILSPVMP 120
 QY 121 YFGLSMLSAISTERCLSIWPIMWYHCRPRYLSSVMCVLWALSILRSILEMFCDFLF 180
 DB 121 YFGLSMLSAISTERCLSIWPIMWYHCRPRYLSSVMCVLWALSILRSILEMFCDFLF 180
 QY 181 SGASVWCEETSDFTIAMLVFLCVLCCSSVLVRLILGSRKMPLTRLYVTILLTVLVF 240
 DB 181 SGASVWCEETSDFTIAMLVFLCVLCCSSVLVRLILGSRKMPLTRLYVTILLTVLVF 240
 QY 241 LILGILPFIQWALFGRHLDKVLFCVHVLVSIIFLSALNSSANPIIYFVGSFRORONQ 300
 DB 241 LILGILPFIQWALFGRHLDKVLFCVHVLVSIIFLSALNSSANPIIYFVGSFRORONQ 300
 QY 301 NLKIVLQRALODTPEVDEGGGWLPOETLELSSRLLEQ 337
 DB 301 NLKIVLQRALODTPEVDEGGGWLPOETLELSSRLLEQ 337

5
 RESULT 2

AA90761
 ID AA90761 standard; protein; 322 AA.
 AC AA90761;
 XX
 XX 18-AUG-2000 (first entry)
 DT
 DE Human G protein-coupled receptor hH17T213 SEQ ID NO:1.
 XX
 XX Human; G protein-coupled receptor; hippocampus; diagnosis; screening;
 KW genetic disease; cellular function regulation.
 XX
 OS Homo sapiens.
 XX
 PN WO200020455-A1.
 XX
 PD 13-APR-2000.
 XX
 PF 30-SEP-1999; 99WO-JP005366.
 XX
 PR 01-OCT-1998; 98JP-00279535.
 XX
 XX (TAKEDA) TAKEDA CHEM IND LTD.
 PA
 XX Watanabe T, Terao Y, Matsui H;
 FI
 DR WPI; 2000-303747/26.
 DR N-PSDB; AAA29811.
 XX

PT Human-derived G protein-coupled protein and encoding nucleic acid, useful
 PT e.g. in determining ligands and treatment of diseases associated with
 PT dysfunction of the protein.
 PS
 Claim 1; Page 90-91; 97pp; Japanese.

CC The present sequence represents a human-derived G protein-coupled protein
 CC designated hH17T213, which is isolated from the human hippocampus. The G
 CC protein-coupled receptor can be used for preventing, treating and
 CC diagnosing genetic diseases associated with G protein-coupled protein,
 CC and for regulating cellular functions. The protein can be used to prevent
 CC and treat disorders associated with G protein-coupled protein gene
 CC dysfunction. It can also be used to identify G protein-coupled protein
 CC ligands and generating antibodies and antisera against the protein. It is
 CC also useful in constructing recombinant receptor protein expression
 CC systems, developing receptor-binding assay systems and screening drug
 CC candidates, and can be used as a probe in the genetic diagnosis of G
 CC protein-coupled protein disorders
 CC
 XX
 SQ Sequence 322 AA;

Query Match 95.7%; Score 1688; DB 3; Length 322;
 Best Local Similarity 100.0%; Pred. No. 1.2e-172;
 Matches 322; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 16 MDSITPVIGTETLPINGREETPCYKQTLSTFGTLCIYSLVALTGNNAVLMILGCRMRNA 75
 DB 1 MDSITPVIGTETLPINGREETPCYKQTLSTFGTLCIYSLVALTGNNAVLMILGCRMRNA 75
 QY 76 VSIYILNVAADFLFSGHIICSPRLINIRHPISKILSPVMPPTPIGSMLSAISTER 135
 DB 76 VSIYILNVAADFLFSGHIICSPRLINIRHPISKILSPVMPPTPIGSMLSAISTER 135
 QY 136 CLSILPMIWHCRPRYLSSVMCVLWALSILRSILEMFCDFLFSGASVWCEETSDFTI 195
 DB 136 CLSILPMIWHCRPRYLSSVMCVLWALSILRSILEMFCDFLFSGASVWCEETSDFTI 195
 QY 195 121 CLSITPMIWHCRPRYLSSVMCVLWALSILRSILEMFCDFLFSGASVWCEETSDFTI 180
 DB 121 CLSITPMIWHCRPRYLSSVMCVLWALSILRSILEMFCDFLFSGASVWCEETSDFTI 180
 QY 255 196 IAMLVFLCVLCCSSVLVRLILGSRKMPLTRLYVTILLTVLVFLCGILPFIQWALFS 255
 DB 196 IAMLVFLCVLCCSSVLVRLILGSRKMPLTRLYVTILLTVLVFLCGILPFIQWALFS 255
 QY 256 RIHLDKVLFCVHVLVSIIFLSALNSSANPIIYFVGSFRORONKLVQRALODTPE 315
 DB 256 RIHLDKVLFCVHVLVSIIFLSALNSSANPIIYFVGSFRORONKLVQRALODTPE 315
 QY 315 241 RIHLDKVLFCVHVLVSIIFLSALNSSANPIIYFVGSFRORONKLVQRALODTPE 300
 DB 241 RIHLDKVLFCVHVLVSIIFLSALNSSANPIIYFVGSFRORONKLVQRALODTPE 300

```
Qy 316 VDEGGWLPQETLELSGSRLQ 337
Db 301 VDEGGWLPQETLELSGSRLQ 322

RESULT 3
ABJ04077
ID ABJ04077 standard; protein; 322 AA.
AC ABJ04077;
AD 11-OCT-2002 (first entry)
DE Human G protein coupled receptor hrup37.
KW Human; G-protein coupled receptor; GPCR; hrup28; hrup29; hrup30; hrup31;
  hrup32; hrup33; hrup34; hrup35; hrup36; hrup37.
XX Homo sapiens.
XX MO200242461-A2.
XX 30-MAY-2002.
XX 26-NOV-2001; 2001MO-US044386.
XX 27-NOV-2000; 2000US-0253404P.
XX 12-DEC-2000; 2000US-0253366P.
XX 20-FEB-2001; 2001US-0270266P.
XX 20-FEB-2001; 2001US-0270266P.
XX 06-APR-2001; 2001US-0282032P.
XX 06-APR-2001; 2001US-0282356P.
XX 06-APR-2001; 2001US-0282358P.
XX 06-APR-2001; 2001US-0282365P.
XX 14-MAY-2001; 2001US-0290917P.
XX 31-JUL-2001; 2001US-0309208P.
XX (AREN-) ARENA PHARM INC.
XX Chen R, Chu ZL, Dang HT, Lowitz KP, Pride C;
  WPI; 2002-566565/60.
XX N-PSDB; ABT04875.
XX Novel endogenous and non-endogenous versions of G protein-coupled
  receptor useful for identification of candidate compounds as receptor
  agonists or antagonists for use as therapeutic agents.
XX PS Claim 37; Page 75-76; 84pp; English.
XX The present invention provides the protein and coding sequences of
  several human G-protein coupled receptors (GPCRs). These can be used in
  the identification of candidate compounds as receptor agonists or inverse
  CC agonists having applicability as therapeutic agents. The present sequence
  CC is a GPCR protein of the invention
XX Sequence 322 AA;

Query Match 95.7%; Score 1688; DB 5; Length 322;
Best Local Similarity 100.0%; Pred. No. 1.2e-172;
Matches 322; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 16 MDSTIPLVLTGTELPINGREBETPCYKQTLSTGLTCTIVSLVATGNNAVVMMLGCRMRRA 75
Db 1 MDSTIPLVLTGTELPINGREBETPCYKQTLSTGLTCTIVSLVATGNNAVVMMLGCRMRRA 60
Qy 76 VSIYIINLVADFLFSGHIIICSPRLINIRHPISTKILSVMTPEPPYIGLSMLSATSTER 135
Db 61 VSIYIINLVADFLFSGHIIICSPRLINIRHPISTKILSVMTPEPPYIGLSMLSATSTER 120
Qy 136 CLSILMPWYHCRPRYLSVSMCVLWALSLNSILEMFCDFLFGSADSVWCETSDPIT 195
  |||
```

```
Db 121 CLSILMPWYHCRPRYLSVSMCVLWALSLNSILEMFCDFLFGSADSVWCETSDPIT 180
Qy 196 IAMLVFLCVVLCGSSLVLRILCGSRKMPRLRYTYTILTVLVFLLCGHPFGIOWALPS 255
Db 181 IAMLVFLCVVLCGSSLVLRILCGSRKMPRLRYTYTILTVLVFLLCGHPFGIOWALPS 240
Qy 256 RIHLDMKVLFCVHVLVSIFLSALNSSANDIYFVGSFRORONRNLKVLGRALQDTPB 315
Db 241 RIHLDMKVLFCVHVLVSIFLSALNSSANDIYFVGSFRORONRNLKVLGRALQDTPB 300
Qy 316 VDEGGWLPQETLELSGSRLQ 337
Db 301 VDEGGWLPQETLELSGSRLQ 322

RESULT 4
AAE21296
ID AAE21296 standard; protein; 322 AA.
XX AAE21296;
XX 01-JUL-2002 (first entry)
XX Human MrgX3 (mas-related gene) protein.
XX Human; mas-related gene; G-protein coupled receptor; drg-12 protein;
  KM receptor; sensory perception; pain; analgesic; MrgX3.
XX Homo sapiens.
XX MO200183555-A2.
XX 08-NOV-2001.
XX 04-MAY-2001; 2001MO-US014519.
XX 04-MAY-2000; 2000US-0202027P.
XX 01-AUG-2000; 2000US-0222344P.
XX 03-NOV-2000; 2000US-00704707.
XX 19-APR-2001; 2001US-0285493P.
XX (CALY) CALIFORNIA INST OF TECHNOLOGY.
XX Anderson DJ, Dong X, Zylka M, Han S, Simon M;
  WPI; 2002-171346/22.
XX N-PSDB; AAD33751.
XX Isolated polypeptide, Mrg, which is a G-protein coupled receptor and an
  PT isolated polypeptide, drg-12, which is also a receptor, useful for
  PT identifying agonists or antagonists for treating pain.
XX PS Claim 16; Page 130; 185pp; English.
XX The invention relates to Mrg (mas-related gene) protein, which is a G-
  CC protein coupled receptor and drg-12 protein, which is a receptor. The
  CC invention is useful for identifying compounds that bind to it, especially
  CC agonists or antagonists. Administration of an agent (e.g. the identified
  CC agonist) that increases the expression of Mrg in a mammal may be used for
  CC treating impaired sensory perception in a mammal, especially pain. The
  CC antagonist may also be useful for treating impaired sensory perception in
  CC a mammal. The present sequence is human MrgX3 protein
XX Sequence 322 AA;

Query Match 95.7%; Score 1688; DB 5; Length 322;
Best Local Similarity 100.0%; Pred. No. 1.2e-172;
Matches 322; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 16 MDSTIPLVLTGTELPINGREBETPCYKQTLSTGLTCTIVSLVATGNNAVVMMLGCRMRRA 75
Db 1 MDSTIPLVLTGTELPINGREBETPCYKQTLSTGLTCTIVSLVATGNNAVVMMLGCRMRRA 60
```


Best Local Similarity 100.0%; Pred. No. 1.2e-172;
Matches 322; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 16 MDSTIPVLTGELTPINREETPCYKQTLSTFTGLCTIYSLALNGNAVVLGCRMRNA 75
Db 1 MDSTIPVLTGELTPINREETPCYKQTLSTFTGLCTIYSLALNGNAVVLGCRMRNA 60

QY 76 VSIYIINLVADFLFSLGHIICSPRLINIRHPISKILSPVMTFPYIGLSMLSAISTER 135
Db 61 VSIYIINLVADFLFSLGHIICSPRLINIRHPISKILSPVMTFPYIGLSMLSAISTER 120

QY 136 CLSILPMIWHCRPRYLSSVMCVLLWALSILSLFEMFCDFLFGSDSVWCETSDFIT 195
Db 121 CLSILPMIWHCRPRYLSSVMCVLLWALSILSLFEMFCDFLFGSDSVWCETSDFIT 180

QY 196 IANLVPLCVLCCSSLVTLVRIILCGSRKMPRLTYTITLTIVLVLCGLPFGIQVALFS 255
Db 181 IANLVPLCVLCCSSLVTLVRIILCGSRKMPRLTYTITLTIVLVLCGLPFGIQVALFS 240

QY 256 RIHDMKVLFCGHVLTSLFSLALSSANPITVFGSFRORONRKLVLQALQDTPR 315
Db 241 RIHDMKVLFCGHVLTSLFSLALSSANPITVFGSFRORONRKLVLQALQDTPR 300

QY 316 VDEGGGWLPOETLELSSGRLEQ 337
Db 301 VDEGGGWLPOETLELSSGRLEQ 322

RESULT 6
AD16998
ID AD16998 standard; protein; 322 AA.
XX AD16998;
AC
XX
DT 15-APR-2004 (first entry)
DE Human NOVX protein homologue Segid 534.
XX
XX human, NOVX; cardiomyopathy; atherosclerosis; cancer; diabetes;
KW inflammation; autoimmune disorder; allergy; blood disorder;
KW acquired immunodeficiency syndrome; AIDS; obesity; asthma;
KW immunoglobulin (Ig)A nephropathy; cirrhosis; arthritis;
KW Alzheimer's disease; infection; str.
XX
OS Homo sapiens.
XX
PN WO200268649-A2.
XX
PD 06-SEP-2002.
XX
XX 31-JAN-2002; 2002MO-US002785.
XX
XX 31-JAN-2001; 2001US-0265395P.
XX 31-JAN-2001; 2001US-0265412P.
XX 31-JAN-2001; 2001US-0265514P.
XX 31-JAN-2001; 2001US-0265517P.
XX 02-FEB-2001; 2001US-0266406P.
XX 05-FEB-2001; 2001US-0266767P.
XX 07-FEB-2001; 2001US-0267057P.
XX 08-FEB-2001; 2001US-0267459P.
XX 09-FEB-2001; 2001US-0267823P.
XX 15-FEB-2001; 2001US-0268974P.
XX 26-FEB-2001; 2001US-0271664P.
XX 27-FEB-2001; 2001US-0271839P.
XX 27-FEB-2001; 2001US-0271855P.
XX 02-MAR-2001; 2001US-0272788P.
XX 02-MAR-2001; 2001US-0273046P.
XX 14-MAR-2001; 2001US-0275925P.
XX 14-MAR-2001; 2001US-0275947P.
XX 14-MAR-2001; 2001US-0275950P.
XX 14-MAR-2001; 2001US-0275989P.
XX 15-MAR-2001; 2001US-0276448P.

PR 15-MAR-2001; 2001US-0276450P.
PR 16-MAR-2001; 2001US-0276397P.
PR 16-MAR-2001; 2001US-0276768P.
PR 20-MAR-2001; 2001US-0278652P.
PR 26-MAR-2001; 2001US-0278775P.
PR 26-MAR-2001; 2001US-0278778P.
PR 26-MAR-2001; 2001US-0279882P.
PR 29-MAR-2001; 2001US-0279884P.
PR 30-MAR-2001; 2001US-0280147P.
PR 11-APR-2001; 2001US-0282992P.
PR 11-APR-2001; 2001US-0283083P.
PR 20-APR-2001; 2001US-0285133P.
PR 23-APR-2001; 2001US-0285749P.
PR 03-MAY-2001; 2001US-0288327P.
PR 03-MAY-2001; 2001US-0288504P.
PR 29-MAY-2001; 2001US-0290447P.
PR 30-MAY-2001; 2001US-0294473P.
PR 08-JUN-2001; 2001US-0296964P.
PR 18-JUN-2001; 2001US-0298959P.
PR 19-JUN-2001; 2001US-0299324P.
PR 13-AUG-2001; 2001US-0312020P.
PR 15-AUG-2001; 2001US-0312889P.
PR 16-AUG-2001; 2001US-0312908P.
PR 21-AUG-2001; 2001US-0313390P.
PR 28-AUG-2001; 2001US-0315470P.
PR 31-AUG-2001; 2001US-0316447P.
PR 07-SEP-2001; 2001US-0318115P.
PR 07-SEP-2001; 2001US-0318118P.
PR 12-SEP-2001; 2001US-0318740P.
PR 19-SEP-2001; 2001US-0323379P.
PR 18-OCT-2001; 2001US-0330245P.
PR 18-OCT-2001; 2001US-0330308P.
PR 14-NOV-2001; 2001US-0332701P.

(CURA-) CURAGEN CORP.
XX
XX Tchernev VT, Spytek KA, Zehruseu BD, Patuaraian M, Shinkets RA;
PI Li L, Gangolli EA, Padgaru M, Anderson DW, Raetelli L, Miller CE;
PI Gerlach VL, Taupier RJ, Gusev VY, Colman SD, Wolenc AR, Pena CRA;
PI Firtak K, Grosse WM, Alsdbrook UF, Lepley DM, Rieger DK, Burgess CE;
XX
XX WPI; 2002-706998/76.
XX
XX New NOVX polypeptides and nucleic acids, useful for preventing or
PT treating NOVX-associated disorders, e.g. cancer, cardiomyopathy,
PT atherosclerosis, or diabetes, and in chromosome mapping, tissue typing or
PT pharmacogenomics.
XX
XX Disclosure; SEQ ID NO 534; 1498bp; English.
XX
XX This invention relates to a novel nucleic acids, and encoded polypeptides
CC thereof, which have properties related to the stimulation of biochemical
CC or physiological responses in a cell, tissue, organ or organism.
CC Specifically, it refers to the use of biologically active fragments for
CC diagnostic and prognostic assays and furthermore in the treatment of
CC diverse pathological conditions. The present invention describes novel
CC human and murine NOVX proteins, as well as methods to modulate their
CC expression using antisense oligos, ribozymes and peptide nucleic acids.
CC The NOVX polypeptides, polynucleotides and antibodies are useful in
CC treating or preventing NOVX-associated disorders, e.g. cardiomyopathy,
CC atherosclerosis, cancer and diabetes. Furthermore, they may be used in
CC treating or preventing diseases such as inflammation, autoimmune
CC disorders, allergies, blood disorders, acquired immunodeficiency syndrome
CC (AIDS), obesity, asthma, immunoglobulin (Ig)A nephropathy, cirrhosis,
CC arthritis, Alzheimer's disease, infections, stroke, muscular dystrophy
CC and epilepsy. Accordingly, these molecules have many activities including
CC cytostatic, cardiant, antiinflammatory, immunosuppressive, antiallergic,
CC hemostatic, anti-HIV, antidiabetic, antiartherosclerotic, anorectic,
CC antisthmatic, nephrotropic, antiarthritic, hepatotropic,
CC neuroprotective, nootropic, antibacterial, virucide, antiparasitic,
CC relaxant and anticonvulsant. In addition, they are useful in screening
CC assays to identify small molecules that modulate or inhibit, for example,
CC neurogenesis, wound healing and angiogenesis. The nucleic acids are also

CC used as in chromosome mapping, tissue typing, preventive medicine and
CC pharmacogenomics. This polypeptide is a homologue of a human NOVX protein
CC of the invention.
XX

SO Sequence 322 AA:

Query Match 95.7%; Score 1688; DB 5; Length 322;
Best Local Similarity 100.0%; Pred. No. 1.2e-172;
Matches 322; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 16 MDSTIPVLTGTELTPTNGREPTPCYKQTLSFTGLTCIVSLVATGNVAVMLGCMRRNA 75
DB 1 MDSTIPVLTGTELTPTNGREPTPCYKQTLSFTGLTCIVSLVATGNVAVMLGCMRRNA 60
QY 76 VSIYILNVADPFLSGHIICSPRLINIRHPISKILSPWTFPYFGLSMLSAISTER 135
DB 61 VSIYILNVADPFLSGHIICSPRLINIRHPISKILSPWTFPYFGLSMLSAISTER 120
QY 136 CLSTIMPWYHCRPRRYLSSVWCVLMLSLRSLIEMWFCDFLFGADSVWCETSDPT 195
DB 121 CLSTIMPWYHCRPRRYLSSVWCVLMLSLRSLIEMWFCDFLFGADSVWCETSDPT 180
QY 196 IAMVFLCVVLCGSSLVLVRLICGSRKMPRLRYVTLITLVVFLCGLPFGIQMALFS 255
DB 181 IAMVFLCVVLCGSSLVLVRLICGSRKMPRLRYVTLITLVVFLCGLPFGIQMALFS 240
QY 256 RIHLDKVLFCFHVHVSIFLSALNSSANPIIYFVGFSPRORONKLVLRALQDTP 315
DB 241 RIHLDKVLFCFHVHVSIFLSALNSSANPIIYFVGFSPRORONKLVLRALQDTP 300
QY 316 VDEGGWLPQETLISGSRLEQ 337
DB 301 VDEGGWLPQETLISGSRLEQ 322

RESULT 7

ABP81750
ID ABP81750 standard; protein; 322 AA.

AC ABP81750;

DT 04-MAR-2003 (first entry)

DE Human G protein-coupled receptor MrgX3 protein SEQ ID NO:674.

XX G protein-coupled receptor; GPCR; antigenic peptide; gene therapy;
XX G protein-coupled receptor modulator; antibody; immune-related disease;
XX growth-related disease; cell regeneration-related disease; AIDS; cancer;
XX immunological-related cell proliferative disease; autoimmune disease;
XX Alzheimer's disease; atherosclerosis; infection; osteoarthritis; allergy;
XX osteoporosis; cardiomyopathy; inflammation; Crohn's disease; diabetes;
XX graft versus host disease; Parkinson's disease; multiple sclerosis; pain;
XX psoriasis; anxiety; depression; schizophrenia; dementia; memory loss;
XX mental retardation; epilepsy; asthma; tuberculosis; obesity; nausea;
XX hyperextension; hypotension; renal disorder; rheumatoid arthritis; trauma;
XX ulcer.

OS Homo sapiens.

XX NO200261087-A2.

XX 08-AUG-2002.

XX 19-DEC-2001; 2001WO-US050107.

XX 19-DEC-2000; 2000US-0257144P.

XX (LIFE-) LIFESPAN BIOSCIENCES INC.

XX Burnet GC, Roush CL, Brown JP;

XX WPI; 2003-046718/04.

XX N-PSDB; ABZ42595.

XX New isolated antigenic peptides e.g., for G protein-coupled receptors
PT (GPCR), useful for diagnosing and designing drugs for treating conditions
PT in which GPCRs are involved, e.g. AIDS, Alzheimer's disease, cancer or
PT autoimmune diseases.

PS Disclosure: Fig 1; 523pp; English.

XX The present invention describes antigenic peptides (1) comprising: (a)
XX any one of 1601 sequences (see ABP82019 to ABP83619) of 12-24 amino
XX acids. Also described: (1) an assay for the detection of a particular G
XX protein-coupled receptor (GPCR) or a candidate polypeptide in a sample;
XX and (2) an isolated antibody having high specificity and high affinity or
XX avidity for a particular GPCR. (1) can be used as GPCR modulators and in
XX gene therapy. The antigenic peptides for GPCRs are useful in detecting an
XX antibody against a particular GPCR, and in the production of specific
XX antibodies. The peptides and antibodies are also useful for detecting the
XX presence or absence of corresponding GPCRs. The antigenic peptides for
XX GPCRs and antibodies are useful for diagnosing and designing drugs for
XX treating immune-related diseases, growth-related diseases, cell
XX regeneration-related disease, immunological-related cell proliferative
XX diseases, or autoimmune diseases, e.g. AIDS, Alzheimer's disease,
XX osteoarthritis, bacterial, fungal, protozoan or viral infections,
XX inflammation, allergies, Crohn's disease, diabetes, graft versus host
XX disease, Parkinson's disease, multiple sclerosis, pain, psoriasis,
XX anxiety, depression, schizophrenia, dementia, mental retardation, memory
XX loss, epilepsy, asthma, tuberculosis, obesity, nausea, hypertension,
XX hypotension, renal disorders, rheumatoid arthritis, trauma, ulcers, or
XX any other disorder in which GPCRs are involved. The antibodies may be
XX used in immunoassays and immunodiagnosis. ABZ42593 to ABZ42869 encode
XX GPCR proteins given in ABP81675 to ABP82018, which are used in the
XX exemplification of the present invention

SO Sequence 322 AA:

Query Match 95.7%; Score 1688; DB 6; Length 322;
Best Local Similarity 100.0%; Pred. No. 1.2e-172;
Matches 322; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 16 MDSTIPVLTGTELTPTNGREPTPCYKQTLSFTGLTCIVSLVATGNVAVMLGCMRRNA 75
DB 1 MDSTIPVLTGTELTPTNGREPTPCYKQTLSFTGLTCIVSLVATGNVAVMLGCMRRNA 60
QY 76 VSIYILNVADPFLSGHIICSPRLINIRHPISKILSPWTFPYFGLSMLSAISTER 135
DB 61 VSIYILNVADPFLSGHIICSPRLINIRHPISKILSPWTFPYFGLSMLSAISTER 120
QY 136 CLSTIMPWYHCRPRRYLSSVWCVLMLSLRSLIEMWFCDFLFGADSVWCETSDPT 195
DB 121 CLSTIMPWYHCRPRRYLSSVWCVLMLSLRSLIEMWFCDFLFGADSVWCETSDPT 180
QY 196 IAMVFLCVVLCGSSLVLVRLICGSRKMPRLRYVTLITLVVFLCGLPFGIQMALFS 255
DB 181 IAMVFLCVVLCGSSLVLVRLICGSRKMPRLRYVTLITLVVFLCGLPFGIQMALFS 240
QY 256 RIHLDKVLFCFHVHVSIFLSALNSSANPIIYFVGFSPRORONKLVLRALQDTP 315
DB 241 RIHLDKVLFCFHVHVSIFLSALNSSANPIIYFVGFSPRORONKLVLRALQDTP 300
QY 316 VDEGGWLPQETLISGSRLEQ 337
DB 301 VDEGGWLPQETLISGSRLEQ 322

RESULT 8

ADH08535
ID ADH08535 standard; protein; 322 AA.

AC ADH08535;

DT 25-MAR-2004 (first entry)

DE MrGX3.
 XX mas-related gene D; MrGD; Analgesic; Vulnerary; Ophthalmological;
 KM sensory perception; glaucoma; MrG.
 XX Mus musculue.
 XX WO2004003133-A1.
 XX 08-JUN-2004.
 XX 13-MAY-2003; 2003MO-US015004.
 XX 26-JUN-2002; 2002US-00183116.
 XX (CALY) CALIFORNIA INST OF TECHNOLOGY.
 XX Anderson DJ, Dong X, Zylka M, Han S, Simon MI;
 XX WPI; 2004-083025/08.
 XX DR N-PSDB; ADH08534.
 XX New mas-related gene D polypeptides, useful as therapeutics or in
 PT identifying agonists or antagonists that alter pain perception in a
 PT mammal for treating impaired sensory perception, e.g. chronic intractable
 PT pain or neuropathic pain.
 XX disclosure, SEQ ID NO 31; 220pp; English.
 XX The present invention relates to an isolated mas-related gene D (MrGD)
 CC polypeptide. The MrGD polypeptides are useful as therapeutics or for
 CC identifying compounds, i.e. agonists or antagonists, that alter pain
 CC perception in a mammal. The compounds are useful for treating impaired
 CC sensory perception, e.g. chronic intractable pain or neuropathic pain,
 CC promoting wound healing, restoring normal sensitivity following injury,
 CC or treating ocular conditions, particularly those associated with
 CC pressure such as glaucoma. The MrG genes or proteins may be used as
 CC molecular probes for the detection of cells or tissues related to or
 CC involved with sensory perception. The present sequence represents a MrGA
 CC (MrG subfamily) protein.
 XX Sequence 322 AA:
 SQ
 Query Match 95.7%; Score 1688; DB 8; Length 322;
 Best Local Similarity 100.0%; Pred. No. 1.2e-172;
 Matches 322; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 16 MDSTIPLVLTGTELPINGREETPCYKOTLSFTGLTCTIVSLVATGNNAVLMLCGRMRNA 75
 DB 1 MDSTIPLVLTGTELPINGREETPCYKOTLSFTGLTCTIVSLVATGNNAVLMLCGRMRNA 60
 QY 76 VSIYIINLVAADEPLFSGHIIICSPRLINIRHPISKILSPVMTFPYFIGLSMISAISTER 135
 DB 61 VSIYIINLVAADEPLFSGHIIICSPRLINIRHPISKILSPVMTFPYFIGLSMISAISTER 120
 QY 136 CISTILPIWTHCRPRYLSSVMCVLMAISLIIEMWPCDPLFGADSVWCETSDFIT 195
 DB 121 CISTILPIWTHCRPRYLSSVMCVLMAISLIIEMWPCDPLFGADSVWCETSDFIT 180
 QY 196 IAMLVFLCVLTCSSSLVLTVRILCGSRKMPLTRLYTILITVAVFLCGPFGIOWALFS 255
 DB 181 IAMLVFLCVLTCSSSLVLTVRILCGSRKMPLTRLYTILITVAVFLCGPFGIOWALFS 240
 QY 256 RIHLDMKVLFCGHVAVSIFLSALNSSANPIIYFVGSFRORONRUKLVQALODTPE 315
 DB 241 RIHLDMKVLFCGHVAVSIFLSALNSSANPIIYFVGSFRORONRUKLVQALODTPE 300
 QY 316 VDEGGGMLPOETLELSSRLQ 337
 DB 301 VDEGGGMLPOETLELSSRLQ 322

AD044602
 ID ADO44602 standard; protein; 322 AA.
 XX ADO44602;
 AC ADO44602;
 XX 29-JUL-2004 (first entry)
 XX Human HIT7213 protein.
 DE HIT7213; transgenic; G protein-coupled receptor; GPCR; ophthalmological;
 KM cytostatic; nephrotoxic; antiinflammatory; dermatological; analgesic;
 KM vulnerary; neuroprotective; human; receptor.
 XX Homo sapiens.
 XX WO2004039972-A1.
 XX 13-MAY-2004.
 XX 28-OCT-2003; 2003MO-JP013781.
 XX 29-OCT-2002; 2002JP-00314141.
 XX (TAKE) TAKEDA CHEM IND LTD.
 XX Katsuo Y, Watanabe T, Yasuhara Y, Mori I, Takeomi S;
 XX WPI; 2004-376191/35.
 XX DR N-PSDB; ADO44603.
 XX HIT7213 protein, encoded DNA and transgenic animals for clarifying
 PT pathological mechanism, developing therapeutic methods and screening
 PT preventives or remedies for related diseases e.g. cataract, cancer, and
 PT dermatitis.
 XX Claim 3; SEQ ID NO 1; 161pp; Japanese.
 XX The invention relates to a non-human mammal that carries a DNA integrated
 CC with a foreign HIT7213 or its mutant gene, or a part of it. The non-human
 CC animal is particularly a rat. Such gene shows phenotypes of e.g. cataract
 CC onset, transient skin rash and proliferation-promoting activity. The
 CC foreign HIT7213 gene is a gene that encodes a G protein-coupled receptor
 CC (GPCR) protein HIT7213. The protein, its encoded DNA and constructed
 CC transgenic animals are useful for clarifying pathological mechanism,
 CC developing therapeutic methods and screening preventives or remedies for
 CC related diseases e.g. cataract, cancer, and dermatitis. The present
 CC sequence represents a human HIT7213 protein.
 XX Sequence 322 AA:
 SQ
 Query Match 95.7%; Score 1688; DB 8; Length 322;
 Best Local Similarity 100.0%; Pred. No. 1.2e-172;
 Matches 322; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 16 MDSTIPLVLTGTELPINGREETPCYKOTLSFTGLTCTIVSLVATGNNAVLMLCGRMRNA 75
 DB 1 MDSTIPLVLTGTELPINGREETPCYKOTLSFTGLTCTIVSLVATGNNAVLMLCGRMRNA 60
 QY 76 VSIYIINLVAADEPLFSGHIIICSPRLINIRHPISKILSPVMTFPYFIGLSMISAISTER 135
 DB 61 VSIYIINLVAADEPLFSGHIIICSPRLINIRHPISKILSPVMTFPYFIGLSMISAISTER 120
 QY 136 CISTILPIWTHCRPRYLSSVMCVLMAISLIIEMWPCDPLFGADSVWCETSDFIT 195
 DB 121 CISTILPIWTHCRPRYLSSVMCVLMAISLIIEMWPCDPLFGADSVWCETSDFIT 180
 QY 196 IAMLVFLCVLTCSSSLVLTVRILCGSRKMPLTRLYTILITVAVFLCGPFGIOWALFS 255
 DB 181 IAMLVFLCVLTCSSSLVLTVRILCGSRKMPLTRLYTILITVAVFLCGPFGIOWALFS 240
 QY 256 RIHLDMKVLFCGHVAVSIFLSALNSSANPIIYFVGSFRORONRUKLVQALODTPE 315
 DB 241 RIHLDMKVLFCGHVAVSIFLSALNSSANPIIYFVGSFRORONRUKLVQALODTPE 300

QY 316 VDEGGWLPQETLEISGRLEQ 337
 DB 301 VDEGGWLPQETLEISGRLEQ 322

RESULT 10
 ADO29705
 ID ADO29705 standard; protein; 322 AA.
 AC ADO29705;
 XX
 DT 29-JUL-2004 (first entry)
 DE Human GPCR MEGX3, SEQ ID NO:807.
 XX
 XX G protein-coupled receptor; GPCR; drug screening; diagnosis;
 KM transgenic mouse; neurological disorder; adrenal gland disorder;
 KM colon disorder; intestinal disorder; cardiovascular disorder;
 KM muscular disorder; blood disorder; immune disorder; bone disorder;
 KM joint disorder; metabolic disorder; nutritive disorder; cancer;
 KM kidney disorder; liver disorder; lung disorder; breast disorder;
 KM ovary disorder; uterus disorder; prostate disorder; testis disorder;
 KM skin disorder; stomach disorder; pancreas disorder; spleen disorder;
 KM thymus disorder; thyroid disorder; antiparkinsonian; antitumor;
 KM cytoskeletal; antiinflammatory; vasotropic; antianginal; antiarrhythmic;
 KM CNS; central nervous system; respiratory; antidiarrhetic; antidiabetic;
 KM virucide; hepatocytic; antibacterial; antianemic; antiseborrhoeic;
 KM dermatological; antitumor; antihypertensive; anorectic;
 KM immunosuppressive; nephrotoxic; gene therapy; GPCR modulator; human;
 KM receptor.
 XX
 OS Homo sapiens.
 XX
 PN MO2004040000-A2.
 XX
 PD 13-MAY-2004.
 XX
 PF 09-SEP-2003; 2003WO-US028226.
 XX
 PR 09-SEP-2002; 2002US-0409303P.
 PR 09-APR-2003; 2003US-0461329P.
 XX
 PA (PRIM-) PRIMAL INC.
 XX
 PI Galtanaris GA, Bergmann JB, Gragerov A, Hohmann J, Li F;
 PI Madisen L, McIlwain KL, Pavlova MN, Vassiliadis D, Zeng H;
 XX
 DR WPI; 2004-399329/36.
 DR N-PSDB; ADO30080.
 XX
 XX Novel mammalian G protein coupled receptors, useful for identifying
 PT compounds that modulates diagnosing and treating disease condition
 PT associated with GPCR dysfunction e.g. autoimmune diseases, angina
 PT pectoris, Parkinson's disease.
 XX
 PS Claim 151; SEQ ID NO 807; 542pp; English.
 XX
 XX The invention relates to human and mouse G protein-coupled receptors
 CC (GPCRs) and nucleic acids encoding them. The invention also relates to
 CC sequences at least 90% identical to the GPCR proteins and nucleic acids
 CC of the invention; methods of treating, preventing or diagnosing diseases
 CC associated with GPCRs of the invention; methods of screening for
 CC compounds useful in the treatment of GPCR-related diseases; a transgenic
 CC mouse comprising a GPCR gene of the invention; a mouse comprising a
 CC mutation in a GPCR transgene or in an endogenous GPCR gene; cells derived
 CC from the transgenic mice; kits comprising several mice, each of which has
 CC a mutation in a different GPCR gene of the invention; and kits comprising
 CC probes which hybridize to GPCR polynucleotides of the invention. The
 CC invention further discloses variants of the GPCR polypeptides and vectors
 CC comprising a GPCR nucleic acid. The GPCR nucleic acids and proteins may
 CC be used in the diagnosis, treatment or prevention of a wide variety of
 CC diseases including neurological disorders (e.g., Alzheimer's disease,

CC depression, diabetic neuropathy, Parkinson's disease or schizophrenia);
 CC disorders of the adrenal gland; disorders of the colon or intestine
 CC (e.g., Crohn's disease, diarrhoea, food poisoning or irritable bowel
 CC syndrome); cardiovascular disorders (e.g., angina, cardiac arrhythmia or
 CC myocardial infarction); muscular disorders; blood disorders (e.g.,
 CC anaemia or leukaemia); immune disorders (e.g., autoimmune disorders or
 CC AIDS); bone and joint disorders (e.g., osteoarthritis, rheumatoid
 CC arthritis, gout or osteoporosis); metabolic or nutritive disorders (e.g.,
 CC obesity, enzyme deficiency-related diseases or vitamin deficiency-related
 CC diseases); and disorders of the kidney, liver, lung, breast, ovary,
 CC uterus, prostate, testis, skin, stomach, pancreas, spleen, thymus and
 CC thyroid (e.g., cancers). The present sequence represents a GPCR of the
 CC invention. Note: The full sequence data for this patent did not form part
 CC of the printed specification; those sequences not shown were obtained in
 CC electronic format directly from WIPO at
 CC ftp.wipo.int/pub/published_pct_sequences.
 XX
 XX Sequence 322 AA;
 SQ

Query Match 95.7%; Score 1688; DB 8; Length 322;
 Best Local Similarity 100.0%; Pred. No. 1,2e-172;
 Matches 322; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 16 MDSTIPVAGTETLPINGREETPCYKQTLSPFGLFCIVSLVATGNAVVLMLGCRMRNA 75
 DB 1 MDSTIPVAGTETLPINGREETPCYKQTLSPFGLFCIVSLVATGNAVVLMLGCRMRNA 60

QY 76 VSIYILNVAADPLFLSGHIIICSPRLINIRHPISKILSPVMTPEYFPGISMLSAISTER 135
 DB 61 VSIYILNVAADPLFLSGHIIICSPRLINIRHPISKILSPVMTPEYFPGISMLSAISTER 120

QY 136 CLSITIPWYHCRPRRYISWVCYLWNLISRLSILEMFCDFLPSGADSVWCETSPFIT 195
 DB 121 CLSITIPWYHCRPRRYISWVCYLWNLISRLSILEMFCDFLPSGADSVWCETSPFIT 180

QY 196 IAMLVFLCVVLGGSLVLLVRIILGSRKMPETRLTYVITLLTVLFLICGLPFGIOMALFS 255
 DB 181 IAMLVFLCVVLGGSLVLLVRIILGSRKMPETRLTYVITLLTVLFLICGLPFGIOMALFS 240

QY 256 RIHLDMKVLFGVHLVSIPLSALNSANPIIYFVGSFRRORONRQNLKVLQRALDPTPE 315
 DB 241 RIHLDMKVLFGVHLVSIPLSALNSANPIIYFVGSFRRORONRQNLKVLQRALDPTPE 300

QY 316 VDEGGWLPQETLEISGRLEQ 337
 DB 301 VDEGGWLPQETLEISGRLEQ 322

RESULT 11
 ADF70481
 ID ADF70481 standard; protein; 560 AA.
 AC ADF70481;
 XX
 DT 12-FEB-2004 (first entry)
 XX
 XX Orphan receptor ligand-related human protein SegID104.
 XX
 XX ligand; orphan receptor protein; fusion protein; fluorescent protein;
 KM cell expression; green fluorescent protein; GFP; GFP-1; wild-type GFP;
 KM GFPuv; Enhanced GFP; EGFP; human.
 XX
 OS Homo sapiens.
 XX
 PN MO2003071272-A1.
 XX
 PD 28-AUG-2003.
 XX
 PF 21-FEB-2003; 2003WO-JP001901.
 XX
 PR 22-FEB-2002; 2002JP-00045728.
 PR 23-JUL-2002; 2002JP-00213949.
 PR 11-OCT-2002; 2002JP-00298237.

XX (TAKE) TAKEDA CHEM IND LTD.
XX
XX Hinuma S, Fujii R, Ogi K, Komatsu H, Kawamata Y, Hosoya M;
XX WPI; 2003-697654/66.
XX DR N-PSDB; ADF70583.
XX
XX Transformation of cells with a fusion protein of an orphan receptor
XX PT protein with a fluorescent protein useful for identification of ligands
XX PT to the orphan receptor.
XX
XX Disclosure; SEQ ID NO 104; 594pp; Japanese.
XX
XX This invention relates to a novel method of identifying ligands to an
XX or orphan receptor protein which comprises transforming cells with DNA
XX encoding a fusion protein of the orphan receptor with a fluorescent
XX protein, so that the fusion protein is expressed in the cells (or cell
XX membranes isolated from them) and contacting the cells with the potential
XX ligand to be tested. A suitable fluorescent protein for incorporation in
XX the fusion protein is green fluorescent protein (GFP), for example GFP-1,
XX wild-type GFP, GFPuv or Enhanced GFP (EGFP). The method is useful for the
XX identification of ligands binding to an orphan receptor protein.
XX
XX Sequence 560 AA:
SQ
Query Match 95.7%; Score 1688; DB 7; Length 560;
Best Local Similarity 100.0%; Pred. No. 2.4e-172;
Matches 322; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 16 MDSTIPVLGTELPIPINGRETPCYKOTLSFTGLTCTIVSLVATGNNAVLMLCGRMRNA 75
DB 1 MDSTIPVLGTELPIPINGRETPCYKOTLSFTGLTCTIVSLVATGNNAVLMLCGRMRNA 60
QY 76 VSIYIINLVADPLFLSGHIICSPRLINIRHPISKILSPVMTFPYFIGLSMLSAISTER 135
DB 61 VSIYIINLVADPLFLSGHIICSPRLINIRHPISKILSPVMTFPYFIGLSMLSAISTER 120
QY 136 CSTIIMPPIWHCRPRYLSSVMCVLMLSLRSIIEMFCDPLFGADSVWCETSPFIT 195
DB 121 CSTIIMPPIWHCRPRYLSSVMCVLMLSLRSIIEMFCDPLFGADSVWCETSPFIT 180
QY 196 IAMLVFLCVLCCSSVLVLRILCGSRKMPLTRLYTITLVAVFLCGLPFGIOWALFS 255
DB 181 IAMLVFLCVLCCSSVLVLRILCGSRKMPLTRLYTITLVAVFLCGLPFGIOWALFS 240
QY 256 RIHLDMKVLFCCHVLVSIPLSALNSSANPIIYFVGSFRORONRUKLVLOALODTPE 315
DB 241 RIHLDMKVLFCCHVLVSIPLSALNSSANPIIYFVGSFRORONRUKLVLOALODTPE 300
QY 316 VDEGGGMLPOETLELSGSRLEQ 337
DB 301 VDEGGGMLPOETLELSGSRLEQ 322
RESULT 12
AA90762
ID AAY90762 standard; protein; 322 AA.
XX
XX AAY90762;
XX
XX 18-AUG-2000 (first entry)
XX
XX Human G protein-coupled receptor hH177213V SEQ ID NO:2.
XX DE
XX Human; G protein-coupled receptor; hippocampus; diagnosis; screening;
XX KM Genetic disease; cellular function regulation.
XX
XX Homo sapiens.
XX
XX WO200020455-A1.
XX
XX 13-APR-2000.
XX PD

XX 30-SEP-1999; 99WO-JP005366.
XX
XX 01-OCT-1998; 98JP-00279535.
XX
XX (TAKE) TAKEDA CHEM IND LTD.
XX
XX Matanabe T, Terao Y, Matsui H;
XX WPI; 2000-303747/26.
XX DR N-PSDB; AAA29812.
XX
XX Human-derived G protein-coupled protein and encoding nucleic acid, useful
XX PT e.g. in determining ligands and treatment of diseases associated with
XX PT dysfunction of the protein.
XX
XX Claim 2; Page 92-93; 97pp; Japanese.
XX
XX The present sequence represents a human-derived G protein-coupled protein
XX designated hH177213V, which is isolated from the human hippocampus. The G
XX protein-coupled receptor can be used for preventing, treating and
XX diagnosing genetic diseases associated with G protein-coupled protein,
XX and for regulating cellular functions. The protein can be used to prevent
XX and treat disorders associated with G protein-coupled protein gene
XX dysfunction. It can also be used to identify G protein-coupled protein
XX ligands and generating antibodies and antisera against the protein. It is
XX also useful in constructing recombinant receptor protein expression
XX systems, developing receptor-binding assay systems and screening drug
XX candidates, and can be used as a probe in the genetic diagnosis of G
XX protein-coupled protein disorders
XX
XX Sequence 322 AA:
SQ
Query Match 95.5%; Score 1683; DB 3; Length 322;
Best Local Similarity 99.7%; Pred. No. 4.1e-172;
Matches 321; Conservative 1; Mismatches 0; Indels 0; Gaps 0;
QY 16 MDSTIPVLGTELPIPINGRETPCYKOTLSFTGLTCTIVSLVATGNNAVLMLCGRMRNA 75
DB 1 MDSTIPVLGTELPIPINGRETPCYKOTLSFTGLTCTIVSLVATGNNAVLMLCGRMRNA 60
QY 76 VSIYIINLVADPLFLSGHIICSPRLINIRHPISKILSPVMTFPYFIGLSMLSAISTER 135
DB 61 VSIYIINLVADPLFLSGHIICSPRLINIRHPISKILSPVMTFPYFIGLSMLSAISTER 120
QY 136 CSTIIMPPIWHCRPRYLSSVMCVLMLSLRSIIEMFCDPLFGADSVWCETSPFIT 195
DB 121 CSTIIMPPIWHCRPRYLSSVMCVLMLSLRSIIEMFCDPLFGADSVWCETSPFIT 180
QY 196 IAMLVFLCVLCCSSVLVLRILCGSRKMPLTRLYTITLVAVFLCGLPFGIOWALFS 255
DB 181 IAMLVFLCVLCCSSVLVLRILCGSRKMPLTRLYTITLVAVFLCGLPFGIOWALFS 240
QY 256 RIHLDMKVLFCCHVLVSIPLSALNSSANPIIYFVGSFRORONRUKLVLOALODTPE 315
DB 241 RIHLDMKVLFCCHVLVSIPLSALNSSANPIIYFVGSFRORONRUKLVLOALODTPE 300
QY 316 VDEGGGMLPOETLELSGSRLEQ 337
DB 301 VDEGGGMLPOETLELSGSRLEQ 322
RESULT 13
AD116994
ID AD116994 standard; protein; 322 AA.
XX
XX AD116994;
XX
XX 15-APR-2004 (first entry)
XX
XX Human NOXV protein homologue SegID 530.
XX DE
XX human; NOXV; cardiomyopathy; atherosclerosis; cancer; diabetes;
XX KM

KM inflammation; autoimmune disorder; allergy; blood disorder
KM acquired immunodeficiency syndrome; AIDS; obesity; asthma;
KM immunoglobulin (Ig) A nephropathy; cirrhosis; arthritis;
KM Alzheimer's disease; infection; str.

OS Homo sapiens.

PN WO200268649-A2.

PD 06-SEP-2002.

PF 31-JAN-2002; 2002WO-US002785.

PR	31-JAN-2001,	2001US-0265195P.
PR	31-JAN-2001,	2001US-0265412P.
PR	31-JAN-2001,	2001US-0265514P.
PR	31-JAN-2001,	2001US-0265517P.
PR	02-FEB-2001,	2001US-0266406P.
PR	05-FEB-2001,	2001US-0266575P.
PR	07-FEB-2001,	2001US-0266975P.
PR	07-FEB-2001,	2001US-0267057P.
PR	08-FEB-2001,	2001US-0267459P.
PR	09-FEB-2001,	2001US-0267823P.
PR	15-FEB-2001,	2001US-0268974P.
PR	26-FEB-2001,	2001US-0271664P.
PR	27-FEB-2001,	2001US-0271839P.
PR	27-FEB-2001,	2001US-0271855P.
PR	02-MAR-2001,	2001US-0272788P.
PR	02-MAR-2001,	2001US-0273046P.
PR	14-MAR-2001,	2001US-0275925P.
PR	14-MAR-2001,	2001US-0275947P.
PR	14-MAR-2001,	2001US-0275950P.
PR	14-MAR-2001,	2001US-0275989P.
PR	15-MAR-2001,	2001US-0276448P.
PR	15-MAR-2001,	2001US-0276450P.
PR	16-MAR-2001,	2001US-0276397P.
PR	16-MAR-2001,	2001US-0276766P.
PR	20-MAR-2001,	2001US-0278652P.
PR	26-MAR-2001,	2001US-0278775P.
PR	26-MAR-2001,	2001US-0278778P.
PR	29-MAR-2001,	2001US-0279882P.
PR	29-MAR-2001,	2001US-0279884P.
PR	30-MAR-2001,	2001US-0280147P.
PR	11-APR-2001,	2001US-0282929P.
PR	11-APR-2001,	2001US-0283083P.
PR	20-APR-2001,	2001US-0285133P.
PR	23-APR-2001,	2001US-0285749P.
PR	03-MAY-2001,	2001US-0288327P.
PR	03-MAY-2001,	2001US-0288504P.
PR	29-MAY-2001,	2001US-0294047P.
PR	30-MAY-2001,	2001US-0294473P.
PR	08-JUN-2001,	2001US-0296964P.
PR	18-JUN-2001,	2001US-0298959P.
PR	19-JUN-2001,	2001US-0299324P.
PR	13-AUG-2001,	2001US-0312020P.
PR	16-AUG-2001,	2001US-0312889P.
PR	16-AUG-2001,	2001US-0312908P.
PR	21-AUG-2001,	2001US-0313390P.
PR	28-AUG-2001,	2001US-0315470P.
PR	31-AUG-2001,	2001US-0316447P.
PR	07-SEP-2001,	2001US-0318115P.
PR	07-SEP-2001,	2001US-0318181P.
PR	12-SEP-2001,	2001US-0318740P.
PR	19-SEP-2001,	2001US-0323379P.
PR	18-OCT-2001,	2001US-0330245P.
PR	18-OCT-2001,	2001US-0330308P.
PR	14-NOV-2001,	2001US-0332701P.

PI Futrak K, Grose WM, Alsbrook JF, Lepley DM, Rieger DK, Burgess CE
 DR WPI: 2002-706938/76.
 XX
 XX New NOVA polypeptides and nucleic acids, useful for preventing or
 PT treating NOVA-associated disorders, e.g. cancer, cardiomyopathy,
 PT atherosclerosis, or diabetes, and in chromosome mapping, tissue typing or
 PT pharmacogenomics.

PS Disclosure; SEQ ID NO 530; 1498pp; English.

This invention relates to a novel nucleic acids, and encoded polypeptides thereof, which have properties related to the stimulation of biochemical or physiological responses in a cell, tissue, organ or organism. Specifically, it refers to the use of biologically active fragments for diagnostic and prognostic assays and furthermore in the treatment of diverse pathological conditions. The present invention describes novel human and murine NOVX proteins, as well as methods to modulate their expression using antisense oligos, ribozymes and peptide nucleic acids. The NOVX polypeptides, polynucleotides and antibodies are useful in treating or preventing NOVX-associated disorders, e.g. cardiomyopathy, atherosclerosis, cancer and diabetes. Furthermore, they may be used in treating or preventing diseases such as inflammation, autoimmune disorders, allergies, blood disorders, acquired immunodeficiency syndrome (AIDS), obesity, asthma, immunoglobulin (Ig)A nephropathy, cirrhosis, arthritis, Alzheimer's disease, infections, stroke, muscular dystrophy and epilepsy. Accordingly, these molecules have many activities including cyostatic, cardiant, antiinflammatory, immunosuppressive, antiallergic, haemostatic, anti-HIV, antidiabetic, antihypertensive, anorectic, antiaesthetic, nephroprotective, antiarthritic, hepatocrotic, neuroprotective, nootropic, antibacterial, vincidine, antiparasitic, relaxant and anticonvulsant. In addition, they are useful in screening assays to identify small molecules that modulate or inhibit, for example, neurogenesis, wound healing and angiogenesis. The nucleic acids are also used as in chromosome mapping, tissue typing, preventive medicine and pharmacogenomics. This polypeptide is a homologue of a human NOVX protein of the invention.

SQ Sequence 322 AA;

Query Match	95.5%	Score 1683;	DB 5;	Length 322;
Best Local Similarity	99.7%	Pred. No. 4.1e-172;		
Matches 321; Conservative	1;	Mismatches 0;	Indels 0;	Gaps 0

Qy	16	MDSTIFPLGTELPI	NGREETPCYQ	YTSFTGLT	ICVSLV	ALTDGA	VYLL	MLGGR	MNA	75
Dd	1	MDSTIFPLGTELPI	NGREBETPCYQ	YTSFTGLT	ICVSLV	ALTDGA	VYLL	MLGGR	MNA	60
Qy	76	VSIIYIINLVAADFL	FLSGHII	ICSP	PLINIRH	PIKSLIP	WMTPEPYI	GLS	MSA	135
Dd	61	VSIIYIINLVAADFL	FLSGHII	ICSP	PLINIRH	PIKSLIP	WMTPEPYI	GLS	MSA	120
Qy	136	CLSLIPIWCHCRPRPYL	SSWMCVLLMAL	SLRSIL	LEMNFC	CDPLFSG	ADSW	MCET	SPIT	195
Dd	121	CLSLIPIWCHCRPRPYL	SSWMCVLLMAL	SLRSIL	LEMNFC	CDPLFSG	ADSW	MCET	SPIT	180
Qy	196	IAMLVPLCVLLCGSS	LVLLVRIL	CGSR	KMP	LR	LYT	ILL	LYLV	255
Dd	181	IAMLVPLCVLLCGSS	LVLLVRIL	CGSR	KMP	LR	LYT	ILL	LYLV	240
Qy	256	RIHLDMPVLCFCHV	LVISIFLS	ALNS	ANP	IYF	FGS	FROR	ONR	315
Dd	241	RIHLDMPVLCFCHV	LVISIFLS	ALNS	ANP	IYF	FGS	FROR	ONR	300
Qy	316	VDEGGGMLPQET	LELSG	SRLEQ						337
Dd	301	VDEGGGMLPQET	LELSG	SRLEQ						322

RESULT 14
AD116999
ID AD116999 standard; protein; 322 AA
XX

AC AD16999;
XX
DT 15-APR-2004 (first entry)
XX
XX Human NOVX protein homologue Segid 535.
XX
XX human; NOVX; cardiomyopathy; atherosclerosis; cancer; diabetes;
XX inflammation; autoimmune disorder; allergy; blood disorder;
XX acquired immunodeficiency syndrome; AIDS; obesity; asthma;
XX immunoglobulin (Ig)A nephropathy; cirrhosis; arthritis;
XX Alzheimer's disease; infection; etc.
XX Homo sapiens.
XX
XX MO200268649-A2.
XX
PD 06-SEP-2002.
XX
XX 31-JAN-2002; 2002MO-US002785.
XX
XX 31-JAN-2001; 2001US-0265395P.
XX 31-JAN-2001; 2001US-0265412P.
XX 31-JAN-2001; 2001US-0265514P.
XX 31-JAN-2001; 2001US-0265517P.
XX 02-FEB-2001; 2001US-0266406P.
XX 05-FEB-2001; 2001US-0266767P.
XX 07-FEB-2001; 2001US-0266975P.
XX 07-FEB-2001; 2001US-0267057P.
XX 08-FEB-2001; 2001US-0267459P.
XX 09-FEB-2001; 2001US-0267823P.
XX 15-FEB-2001; 2001US-0268974P.
XX 26-FEB-2001; 2001US-0271664P.
XX 27-FEB-2001; 2001US-0271839P.
XX 27-FEB-2001; 2001US-0271855P.
XX 02-MAR-2001; 2001US-0272788P.
XX 02-MAR-2001; 2001US-0273046P.
XX 14-MAR-2001; 2001US-0275925P.
XX 14-MAR-2001; 2001US-0275947P.
XX 14-MAR-2001; 2001US-0275950P.
XX 14-MAR-2001; 2001US-0275989P.
XX 15-MAR-2001; 2001US-0276448P.
XX 15-MAR-2001; 2001US-0276450P.
XX 16-MAR-2001; 2001US-0276397P.
XX 16-MAR-2001; 2001US-0276768P.
XX 20-MAR-2001; 2001US-0278652P.
XX 26-MAR-2001; 2001US-0278775P.
XX 26-MAR-2001; 2001US-0278778P.
XX 29-MAR-2001; 2001US-0279882P.
XX 29-MAR-2001; 2001US-0279884P.
XX 30-MAR-2001; 2001US-0280147P.
XX 11-APR-2001; 2001US-0282992P.
XX 11-APR-2001; 2001US-0283083P.
XX 20-APR-2001; 2001US-0285133P.
XX 23-APR-2001; 2001US-0285749P.
XX 03-MAY-2001; 2001US-0288327P.
XX 03-MAY-2001; 2001US-0288504P.
XX 29-MAY-2001; 2001US-0294047P.
XX 30-MAY-2001; 2001US-0294473P.
XX 08-JUN-2001; 2001US-0296964P.
XX 18-JUN-2001; 2001US-0298959P.
XX 19-JUN-2001; 2001US-0299324P.
XX 13-AUG-2001; 2001US-0312020P.
XX 16-AUG-2001; 2001US-0312889P.
XX 16-AUG-2001; 2001US-0312908P.
XX 21-AUG-2001; 2001US-0313390P.
XX 28-AUG-2001; 2001US-0315470P.
XX 31-AUG-2001; 2001US-0316447P.
XX 07-SEP-2001; 2001US-0318115P.
XX 07-SEP-2001; 2001US-0318118P.
XX 12-SEP-2001; 2001US-0318740P.
XX 19-SEP-2001; 2001US-0323379P.
XX 18-OCT-2001; 2001US-0330245P.
XX 18-OCT-2001; 2001US-0330308P.

PR 14-NOV-2001; 2001US-0332701P.
XX
XX (CURA-) CURAGEN CORP.
XX
XX Tchernev VT, Spytek KA, Zernhusen BD, Paturnajan M, Shinkens RA;
PI Li L, Gangoli EA, Padigaru M, Anderson DW, Raestelli L, Miller CB;
PI Gerlach VL, Taupier RJ, Gusev VV, Colman SD, Wolenc AR, Pena CEA; CB;
PI Furtak K, Grosse WM, Alsbrook JF, Lepley DM, Rieger DK, Burgess CB;
XX
XX WPI; 2002-706998/76.
XX
XX
XX New NOVX polypeptides and nucleic acids, useful for preventing or
PT treating NOVX-associated disorders, e.g. cancer, cardiomyopathy,
PT atherosclerosis, or diabetes, and in chromosome mapping, tissue typing or
PT pharmacogenomics.
XX
XX
XX Disclosure, SEQ ID NO 535; 1498bp; English.
XX
XX
XX This invention relates to a novel nucleic acids, and encoded polypeptides
CC thereof, which have properties related to the stimulation of biochemical
CC or physiological responses in a cell, tissue, organ or organism.
CC Specifically, it refers to the use of biologically active fragments for
CC diagnostic and prognostic assays and furthermore in the treatment of
CC diverse pathological conditions. The present invention describes novel
CC human and murine NOVX proteins, as well as methods to modulate their
CC expression using antisense oligos, ribozymes and peptide nucleic acids.
CC The NOVX polypeptides, polynucleotides and antibodies are useful in
CC treating or preventing NOVX-associated disorders, e.g. cardiomyopathy,
CC atherosclerosis, cancer and diabetes. Furthermore, they may be used in
CC treating or preventing diseases such as inflammation, autoimmune
CC disorders, allergies, blood disorders, acquired immunodeficiency syndrome
CC (AIDS), obesity, asthma, immunoglobulin (Ig)A nephropathy, cirrhosis,
CC arthritis, Alzheimer's disease, infections, stroke, muscular dystrophy
CC and epilepsy. Accordingly, these molecules have many activities including
CC cytostatic, cardiant, anti-inflammatory, immunosuppressive, antiallergic,
CC haemostatic, anti-HIV, antidiabetic, antiatherosclerotic, anorectic,
CC antistomatitic, nephrotropic, anticholelithic, hepatotropic,
CC neuroprotective, nootropic, antibacterial, virucide, antiparasitic,
CC relaxant and anticonvulsant. In addition, they are useful in screening
CC assays to identify small molecules that modulate or inhibit, for example,
CC neurogenesis, wound healing and angiogenesis. The nucleic acids are also
CC used as in chromosome mapping, tissue typing, preventive medicine and
CC pharmacogenomics. This polypeptide is a homologue of a human NOVX protein
CC of the invention.
XX
XX
XX Sequence 322 AA;
SQ
Query Match 95.5%; Score 1683; DB 5; Length 322;
Best Local Similarity 99.7%; Pred. No. 4.1e-172;
Matches 321; Conservative 1; Mismatches 0; Indels 0; Gaps 0;
QY 16 MSTIPLYLGTLPINGRETPCYKOTLSPTGTCIVSLVALTGNVVMMLGCRMRNA 75
DB 1 MSTIPLYLGTLPINGRETPCYKOTLSPTGTCIVSLVALTGNVVMMLGCRMRNA 60
QY 76 VSIYIINLVAAADFLPSGHICSPRLINIRHPSIKILSPVMTFPFIFGISMLSAISTER 135
DB 61 VSIYIINLVAAADFLPSGHICSPRLINIRHPSIKILSPVMTFPFIFGISMLSAISTER 120
QY 136 CSLIWPIMWYHCRPRYLSVWCVLMLALSLSILEMFCDFLFGSGADSVWCETSDFIT 195
DB 121 CSLIWPIMWYHCRPRYLSVWCVLMLALSLSILEMFCDFLFGSGADSVWCETSDFIT 180
QY 196 IAWLVFLCVVLCSSVLVLRILICGRKMPRLRYTITLLTVLVFLCGIPRIGIOWALFS 255
DB 181 IAWLVFLCVVLCSSVLVLRILICGRKMPRLRYTITLLTVLVFLCGIPRIGIOWALFS 240
QY 256 RIHLDMKVLFCVHVLVIFLSALNSSANPIYFVGSFPROKORNLKVLQBALDPTFS 315
DB 241 RIHLDMKVLFCVHVLVIFLSALNSSANPIYFVGSFPROKORNLKVLQBALDPTFS 300
QY 316 VDEGGMLPQETLLESGRLAQ 337

Db 301 VDEGGWLPQETLBSGSRLQ 322

RESULT 15

ID ADC86821 standard; protein, 322 AA.

AC ADC86821;

DT 01-JAN-2004 (first entry)

DE Human GPCR protein SEQ ID NO:1274.

human; GPCR; guanosine triphosphate-binding protein coupled receptor; gene therapy.

OS Homo sapiens.

PN EPI270724-A2.

PD 02-JAN-2003.

PF 18-JUN-2002; 2002EP-00013517.

PR 18-JUN-2001; 2001JP-00246789.

PA (NAAD-) NAT INST ADVANCED IND SCI & TECHNOLOGY.
PA (ADSC-) CENT ADVANCED SCI & TECHNOLOGY INCUBATIO.

PI Suwa M, Asai K, Akiyama Y, Aburatani H;

DR WPI; 2003-315783/31.

DR N-PSDB; ADC86820.

PT New polynucleotide, useful for preparing a composition for treating a patient in need of increased or suppressed activity or expression of the guanosine triphosphate-binding protein coupled receptor.

PS Claim 2; SEQ ID NO 1274; 28bp; English.

CC The invention relates to a novel polynucleotide encoding a guanosine triphosphate-binding protein coupled receptor (GPCR). A polynucleotide of the invention may have a use in gene therapy. The polynucleotide and polypeptide are useful for preparing a composition for treating a patient in need of increased or suppressed activity or expression of the guanosine triphosphate-binding protein coupled receptor. The protein sequences shown in ADC85549-ADC87617 represent GPCR's of the invention.

XX Sequence 322 AA;

Query Match 95.5%; Score 1683; DB 7; Length 322;

Best Local Similarity 99.7%; Pred. No. 4,1e-172;

Matches 321; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

16 MDSTIPVLGTELTPINGREETPCYKQTLSPFTGLTCTIVSLVATGNAVVLMLGCMRRNA 75

1 MDSTIPVLGTELTPINGREETPCYKQTLSPFTGLTCTIVSLVATGNAVVLMLGCMRRNA 60

76 VSIYILNVADPLFLSGHIIICSPRLINIRHPISKIISPVTTPPYFTGLSMLSAISTER 135

61 VSIYILNVADPLFLSGHIIICSPRLINIRHPISKIISPVTTPPYFTGLSMLSAISTER 120

136 CLSTLMPITWYHCRPRYSVYVLCVLLMALSLRLSILEMWFCDPLFSGADSVWCETSDFTT 195

121 CLSTLMPITWYHCRPRYSVYVLCVLLMALSLRLSILEMWFCDPLFSGADSVWCETSDFTT 180

196 IAWLVFLCVLTCGSSLVLLVRLICGSRKMPLTRLYVTILLTVLVFLCGLPFGIQWALFS 255

181 IAWLVFLCVLTCGSSLVLLVRLICGSRKMPLTRLYVTILLTVLVFLCGLPFGIQWALFS 240

256 RIHIDMKVLFCHVHVSIFLSALNSSANPIIYFFVGSFRQRONRNLKLVLRALQDTPR 315

241 RIHIDMKVLFCHVHVSIFLSALNSSANPIIYFFVGSFRQRONRNLKLVLRALQDTPR 300

Qy 316 VDEGGWLPQETLBSGSRLQ 337

Db 301 VDEGGWLPQETLBSGSRLQ 322

Search completed: October 27, 2004, 08:30:17
Job time : 160 secs

GenCore version 5.1.6
Copyright (c) 1993 - 2004 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: October 27, 2004, 08:24:28 ; Search time 40 Seconds
(without alignments)
558.729 Million cell updates/sec

Title: US-09-867-570-2

Perfect score: 1763
Sequence: 1 MSKSSWVIRLGFSLSDSTI.....EGGWLPGFTLSSRLSQ 337

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 478139 seqs, 6631800 residues

Total number of hits satisfying chosen parameters: 478139

Minimum DB seq length: 0
Maximum DB seq length: 200000000

Post-processing: Minimum Match 0%
Listing first 45 summaries

Database : Issued Patents AA:*
1: /cgn2_6/prodata/1/aa/5A COMB pep:*
2: /cgn2_6/prodata/1/aa/5B COMB pep:*
3: /cgn2_6/prodata/1/aa/6A COMB pep:*
4: /cgn2_6/prodata/1/aa/6B COMB pep:*
5: /cgn2_6/prodata/1/aa/PCTUS COMB pep:*
6: /cgn2_6/prodata/1/aa/backfile1.pep:*

Pred. No. is the number of results predicted by chance to have a
score greater than or equal to the score of the result being printed,
and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	1642	93.1	322	4	US-09-254-227A-3 Sequence 3, Appl1
2	1598	90.6	322	4	US-09-254-227A-5 Sequence 5, Appl1
3	1405	79.7	322	4	US-09-254-227A-7 Sequence 7, Appl1
4	1391	78.9	322	4	US-09-254-227A-9 Sequence 9, Appl1
5	1379	78.2	322	4	US-09-254-227A-11 Sequence 11, Appl1
6	1360	77.1	322	4	US-09-254-227A-13 Sequence 13, Appl1
7	834.5	47.3	337	4	US-09-254-227A-1 Sequence 1, Appl1
8	423.5	24.0	325	6	5320941-2 Patent No. 5320941
9	388.5	22.0	282	1	US-08-118-270-52 Sequence 52, Appl1
10	388.5	22.0	282	5	PCT-US93-08528-52 Sequence 52, Appl1
11	344	19.5	298	1	US-08-118-270-76 Sequence 76, Appl1
12	344	19.5	298	5	PCT-US93-08528-76 Sequence 76, Appl1
13	255.5	14.5	335	3	US-08-981-825-6 Sequence 6, Appl1
14	255.5	14.5	335	3	US-09-480-784-6 Sequence 6, Appl1
15	235.5	13.4	354	3	US-07-759-568-2 Sequence 2, Appl1
16	230.5	13.1	381	4	US-09-745-842-21 Sequence 21, Appl1
17	226.5	12.8	355	1	US-07-759-568-1 Sequence 1, Appl1
18	226.5	12.8	355	1	US-08-450-393A-8 Sequence 8, Appl1
19	226.5	12.8	355	2	US-08-390-000A-5 Sequence 5, Appl1
20	226.5	12.8	355	3	US-08-446-669-8 Sequence 8, Appl1
21	226.5	12.8	355	4	US-09-625-573-8 Sequence 8, Appl1
22	226.5	12.8	355	5	PCT-US95-00476-8 Sequence 8, Appl1
23	226.5	12.8	350	1	US-08-202-056-7 Sequence 7, Appl1
24	226.5	12.8	360	4	US-09-409-778-4 Sequence 4, Appl1
25	219.5	12.5	369	1	US-07-816-283-8 Sequence 8, Appl1
26	219.5	12.5	369	1	US-08-417-103-8 Sequence 8, Appl1
27	219.5	12.5	369	2	US-08-411-859-3 Sequence 3, Appl1

28	219.5	12.5	369	3	US-08-120-601B-9 Sequence 9, Appl1
29	219.5	12.5	369	4	US-08-387-707-9 Sequence 9, Appl1
30	219.5	12.5	369	4	US-08-405-271A-9 Sequence 9, Appl1
31	218.5	12.4	381	1	US-08-467-125-2 Sequence 2, Appl1
32	218.5	12.4	381	2	US-08-911-320A-2 Sequence 2, Appl1
33	218.5	12.4	381	2	US-09-217-101-2 Sequence 2, Appl1
34	217	12.3	351	4	US-09-944-807-2 Sequence 2, Appl1
35	211	12.0	353	4	US-09-576-160B-6 Sequence 6, Appl1
36	210	11.9	259	4	US-09-261-599B-3 Sequence 3, Appl1
37	210	11.9	259	4	US-09-456-455A-3 Sequence 3, Appl1
38	210	11.9	380	3	US-08-676-351-5 Sequence 5, Appl1
39	206.5	11.7	384	3	US-09-071-434-3 Sequence 5, Appl1
40	206	11.7	380	3	US-08-188-275A-5 Sequence 5, Appl1
41	206	11.7	380	3	US-09-351-198-5 Sequence 5, Appl1
42	206	11.7	380	3	US-09-113-426-5 Sequence 5, Appl1
43	205.5	11.7	355	4	US-09-170-496D-2 Sequence 2, Appl1
44	205	11.6	333	4	US-09-170-496D-8 Sequence 8, Appl1
45	204	11.6	333	4	US-09-170-496D-168 Sequence 168, App

ALIGNMENTS

RESULT 1	US-09-254-227A-3	US-09-254-227A-3
Sequence 3, Application US/09254227A		
Patent No. 6696257		
GENERAL INFORMATION:		
APPLICANT: Amad, Sultan		
APPLICANT: Banville, Denis		
APPLICANT: Fortin, Yves		
APPLICANT: Lembo, Paola		
APPLICANT: O'Donnell, Dajan		
APPLICANT: Shi-Hsiang, Shen		
TITLE OF INVENTION: G Protein-Coupled Receptors from the Rat and Human		
FILE REFERENCE: 81823/268117		
CURRENT APPLICATION NUMBER: US/09/254,227A		
CURRENT FILING DATE: 1999-03-03		
NUMBER OF SEQ ID NOS: 22		
SOFTWARE: Patent version 3.0		
SEQ ID NO 3		
LENGTH: 322		
TYPE: PRT		
ORGANISM: Homo sapiens		
US-09-254-227A-3		
Query Match	93.1%	Score 1642; DB 4; Length 322;
Best Local Similarity	97.2%;	Pred. No. 8.7e-138;
Matches 313; Conservative	5; Mismatches	4; Indels 0; Gaps 0;
QY	16	MDSTIVLGTSLTPINGRETPCYKOTLSFTGTCIVSLVATLGNVLMGLGCRMRNA 75
DB	1	MDSTIVLGTSLTPINGRETPCYKOTLSFTGTCIVSLVATLGNVLMGLGCRMRNA 60
QY	76	VSIIYINLVADPLFLSGHICSPRLINIRHPIKSLSPVMTFPYFISGLMSAISTER 135
DB	61	VSIIYINLVADPLFLSGHICSPRLINIRHPIKSLSPVMTFPYFISGLMSAISTER 120
QY	136	CSILMPWYHCRPRYLSVWCVLLMALSLRSIEMFCDPLFGSADSWCETSDFIT 195
DB	121	CSILMPWYHCRPRYLSVWCVLLMALSLRSIEMFCDPLFGSADSWCETSDFIT 180
QY	196	IAMLVLCVYLCCSSLVLRILICGRKMPRLTYTILITVAVFLCGIPRIGQALFS 255
DB	181	IAMLVLCVYLCCSSLVLRILICGRKMPRLTYTILITVAVFLCGIPRIGQALFS 240
QY	256	RHLDMKVLFCVHLVSIPLSALNSSANPIYFVGSFROKORONLKVLOALDTPR 315
DB	241	RHLDMKVLFCVHLVSIPLSALNSSANPIYFVGSFROKORONLKVLOALDTPR 300
QY	316	VDEGGWLPQETLSSRLSQ 337
DB	301	VDEGGWLPQETLSSRLSQ 322

```
RESULT 2
US-09-254-227A-5
; Sequence 5, Application US/09254227A
; Patent No. 6696257
; GENERAL INFORMATION:
; APPLICANT: Ahmad, Sultan
; APPLICANT: Banville, Denis
; APPLICANT: Fortin, Yves
; APPLICANT: Lembo, Paola
; APPLICANT: O'Donnell, Dajan
; APPLICANT: Shi-Hsiang, Shen
; TITLE OF INVENTION: G Protein-Coupled Receptors from the Rat and Human
; FILE REFERENCE: 81823/268117
; CURRENT APPLICATION NUMBER: US/09/254,227A
; NUMBER OF SEQ ID NOS: 22
; SOFTWARE: Patentin version 3.0
; SEQ ID NO 5
; LENGTH: 322
; TYPE: PRT
; ORGANISM: Homo sapiens
US-09-254-227A-5

Query Match          90.6%; Score 1598; DB 4; Length 322;
Best Local Similarity 95.6%; Pred. No. 6.9e-134;
Matches 307; Conservative 5; Mismatches 9; Indels 0; Gaps 0;

QY 16 MDSTIPVLGTELTPTINGREETPCYKQTLSFTGLCTIYSLVATLGNNAVLMILGCMRRNA 75
DB 1 MDPTVPVLGTELTPTINGREETPCYKQTLSFTGLCTIYSLVATLGNNAVLMILGCMRRNA 60

QY 76 VSIYIILNVAADFLPLSGHIICSPRLINIRHPISKILSPWTFPPYFGLSMISAISTER 135
DB 1 MDPTVPVLGTELTPTINGREETPCYKQTLSFTGLCTIYSLVATLGNNAVLMILGCMRRNA 60

QY 136 CLSLMPDWYHCRPRYLSSVMCVLWALSLRSILEMFCDFLFGSADSVWCETSDPT 195
DB 61 VSIYIILNVAADFLPLSGHIICSPRLINIRHPISKILSPWTFPPYFGLSMISAISTER 120

QY 121 CLSLMPDWYHCRPRYLSSVMCVLWALSLRSILEMFCDFLFGSADSVWCETSDPT 180
DB 121 CLSLMPDWYHCRPRYLSSVMCVLWALSLRSILEMFCDFLFGSADSVWCETSDPT 180

QY 196 IAWLVFLCVLGGSSVLVLRILGSRKMPRLRYVTIILTVLVFLGSPFGIOWALFS 255
DB 181 IAWLVFLCVLGGSSVLVLRILGSRKMPRLRYVTIILTVLVFLGSPFGIOWALFS 240

QY 256 RIHLDWKVLFCHVALVSIPLSALNSSANPIIYFVGSFRORONRKLVLQALQDTP 315
DB 241 RIHLDWKVLFCHVALVSIPLSALNSSANPIIYFVGSFRORONRKLVLQALQDTP 300

QY 316 VDEGGMWLPQETTELSSGRLE 336
DB 301 VDEGGMWLPQETTELSSGRLE 321

RESULT 3
US-09-254-227A-7
; Sequence 7, Application US/09254227A
; Patent No. 6696257
; GENERAL INFORMATION:
; APPLICANT: Ahmad, Sultan
; APPLICANT: Banville, Denis
; APPLICANT: Fortin, Yves
; APPLICANT: Lembo, Paola
; APPLICANT: O'Donnell, Dajan
; APPLICANT: Shi-Hsiang, Shen
; TITLE OF INVENTION: G Protein-Coupled Receptors from the Rat and Human
; FILE REFERENCE: 81823/268117
; CURRENT APPLICATION NUMBER: US/09/254,227A
; NUMBER OF SEQ ID NOS: 22
; SOFTWARE: Patentin version 3.0
; SEQ ID NO 7
; LENGTH: 322
```

```
; TYPE: PRT
; ORGANISM: Homo sapiens
US-09-254-227A-7

Query Match          79.7%; Score 1405; DB 4; Length 322;
Best Local Similarity 83.9%; Pred. No. 8.7e-117;
Matches 270; Conservative 18; Mismatches 34; Indels 0; Gaps 0;

QY 16 MDSTIPVLGTELTPTINGREETPCYKQTLSFTGLCTIYSLVATLGNNAVLMILGCMRRNA 75
DB 1 MDPTVPVLGTELTPTINGREETPCYKQTLSFTGLCTIYSLVATLGNNAVLMILGCMRRNA 60

QY 76 VSIYIILNVAADFLPLSGHIICSPRLINIRHPISKILSPWTFPPYFGLSMISAISTER 135
DB 61 VSIYIILNVAADFLPLSGHIICSPRLINIRHPISKILSPWTFPPYFGLSMISAISTER 120

QY 136 CLSLMPDWYHCRPRYLSSVMCVLWALSLRSILEMFCDFLFGSADSVWCETSDPT 195
DB 121 CLSLMPDWYHCRPRYLSSVMCVLWALSLRSILEMFCDFLFGSADSVWCETSDPT 180

QY 196 IAWLVFLCVLGGSSVLVLRILGSRKMPRLRYVTIILTVLVFLGSPFGIOWALFS 255
DB 181 IAWLVFLCVLGGSSVLVLRILGSRKMPRLRYVTIILTVLVFLGSPFGIOWALFS 240

QY 256 RIHLDWKVLFCHVALVSIPLSALNSSANPIIYFVGSFRORONRKLVLQALQDTP 315
DB 241 RIHLDWKVLFCHVALVSIPLSALNSSANPIIYFVGSFRORONRKLVLQALQDTP 300

QY 316 VDEGGMWLPQETTELSSGRLEQ 337
DB 301 VDEGGMWLPQETTELSSGRLEQ 322
```

```
RESULT 4
US-09-254-227A-9
; Sequence 9, Application US/09254227A
; Patent No. 6696257
; GENERAL INFORMATION:
; APPLICANT: Ahmad, Sultan
; APPLICANT: Banville, Denis
; APPLICANT: Fortin, Yves
; APPLICANT: Lembo, Paola
; APPLICANT: O'Donnell, Dajan
; APPLICANT: Shi-Hsiang, Shen
; TITLE OF INVENTION: G Protein-Coupled Receptors from the Rat and Human
; FILE REFERENCE: 81823/268117
; CURRENT APPLICATION NUMBER: US/09/254,227A
; NUMBER OF SEQ ID NOS: 22
; SOFTWARE: Patentin version 3.0
; SEQ ID NO 9
; LENGTH: 322
; TYPE: PRT
; ORGANISM: Homo sapiens
US-09-254-227A-9

Query Match          78.9%; Score 1391; DB 4; Length 322;
Best Local Similarity 83.9%; Pred. No. 1.5e-115;
Matches 270; Conservative 16; Mismatches 36; Indels 0; Gaps 0;

QY 16 MDSTIPVLGTELTPTINGREETPCYKQTLSFTGLCTIYSLVATLGNNAVLMILGCMRRNA 75
DB 1 MDPTVPVLGTELTPTINGREETPCYKQTLSFTGLCTIYSLVATLGNNAVLMILGCMRRNA 60

QY 76 VSIYIILNVAADFLPLSGHIICSPRLINIRHPISKILSPWTFPPYFGLSMISAISTER 135
DB 61 VSIYIILNVAADFLPLSGHIICSPRLINIRHPISKILSPWTFPPYFGLSMISAISTER 120

QY 136 CLSLMPDWYHCRPRYLSSVMCVLWALSLRSILEMFCDFLFGSADSVWCETSDPT 195
DB 121 CLSLMPDWYHCRPRYLSSVMCVLWALSLRSILEMFCDFLFGSADSVWCETSDPT 180

QY 196 IAWLVFLCVLGGSSVLVLRILGSRKMPRLRYVTIILTVLVFLGSPFGIOWALFS 255
```

Db 181 VAMVILFCLVVLCCSSLYVLRILKCSRKILPLTRLYVTLLVYVFLGCLPFGIQFPLFL 240
QY 256 RHLDMKVLFCVHVLVSIPLSLANSSANPIIYFVGSFRORONIKLVIQRALDTPK 315
Db 241 WIHVDEEVLFCHVHLVSIPLSLANSSANPIIYFVGSFRORONIKLVIQRALDASE 300
QY 316 VDEGGGMLPQETTELGSRLQ 337
Db 301 VDEGGGQLPQETTELGSRLQ 322

```

RESULT 5
US-09-254-227A-11
: Sequence 11, Application US/09254227A
: Patent No. 6696257
: GENERAL INFORMATION:
: APPLICANT: Ahmad, Sultan
: APPLICANT: Banville, Denis
: APPLICANT: Portin, Yves
: APPLICANT: Lembo, Paola
: APPLICANT: O'Donnell, Dajan
: APPLICANT: Shi-Hsiang, Shen
: TITLE OF INVENTION: G Protein-Coupled Receptors from the Rat and Human
: FILE REFERENCE: 81823/268117
: CURRENT FILING DATE: 1999-03-03
: NUMBER OF SEQ ID NOS: 22
: SOFTWARE: PatentIn version 3.0
: SEQ ID NO 11
: LENGTH: 322
: TYPE: PRS
: ORGANISM: Homo sapiens
: US-09-254-227A-11

```

Query Match	78.2%	Score 1379	DB 4	Length 332
Best Local Similarity	83.4%	Pred. No. 1.8e-114		
Matches 267	Conservative 20	Mismatches 33	Indels 0	Gaps 0

Qy	16	MSSTIIVLSTELTLPINGRETEPCYQOTLSFTGLCTIVALTGNAVLMILGGRMRNA	75
Db	1	MDPTAVLSTKGLPIINGRETEPCYQOTLSFTGLCTISLVGIGTNAVVLMLGGRMRNA	60
Qy	76	VSIIYIINLVADPLFLISGHIICSPRLINIRIPISKLSPWMTTPPYFLGSLMSAISTER	135
Db	61	VSIIYIINLVADPLFLISFQIICRPRLINISHLIKRIIVSWTPPYFLGSLMSAISTER	120
Qy	136	CLSIILMPIYHCRPRPYLSSVMCVLLMALSLRSILEMFCDFPLSGADSVWCETSDPIT	195
Db	121	CLSIIVMPIYRCRPRPHLSAAVVCVLLMAGLLFLSLMEHRFCDFPLSGADSVWCETSDPIT	180
Qy	196	IAMLVPLCYVLGCGSSLVTLVRLICGRKMPRLRYTITLLTVLVFLCGLPFGIQLMFLS	255
Db	181	VAMLVPLCYVLGCGSSLVTLVRLICGRKMPRLRYTITLLTVLVFLCGLPFGIQLMFLY	240
Qy	256	RILHDKVPLFCCHLVLSIFSLAINSSANPIIYFPGSFROKRONKTLVQRLADTPPE	315
Db	241	RHILHDKVPLFCCHLVLYCMSELSSAINSANPIIYFPGSFROKRONKTLVQRLADTPPE	300
Qy	316	VDEGGGMLPOETLTLSGSRLL	335
Db	301	VDEGGGMLPOETLTLSGSRLL	320

RESULT 6
US-09-254-227A-13
Sequence 13, Application US/09254227A
Parent No. 6696757
GENERAL INFORMATION:
APPLICANT: Ahmad, Sultan
APPLICANT: Barville, Denis
APPLICANT: Fortin, Yves
APPLICANT: Lembo, Paola

```

; APPLICANT: O'Donnell, Dajan
; APPLICANT: Shi-Hsiang, Shen
; TITLE OF INVENTION: G Protein-Coupled Receptors from the Rat and Human
; FILE REFERENCE: 81823/268117
; CURRENT APPLICATION NUMBER: US/09/254,227A
; CURRENT FILING DATE: 1999-03-03
; NUMBER OF SEQ ID NOS: 22
; SOFTWARE: PatentIn version 3.0
; SEQ ID NO 13
; LENGTH: 322
; TYPE: PRT
; ORGANISM: Homo sapiens
; US-09-254-227A-13

```

Query Match	77.1%	Score 1360;	DB 4;	Length 322;
Best Local Similarity	82.5%	Pred. No. 8	5e-113;	
Matches 264;	Conservative 21;	Mismatches 35;	Indels 0;	Gaps 0;

Qy	16	MSSTIVGATETLTPINGREBETPCYQQTSPFGTICIVSLVMTNNAVLTMLGGKMRNA	75
Db	1	MDPTAVFGFKLTPINGREBETPCYQQTSPFVLTCTIISLVGTGNAAVLTMLGYRMRNA	60
Qy	76	VSYIYIINAAAPLFLSGHITCSPTLRLNIRHPTLSKILSPVMTPEPYPLGISMISAISTER	135
Db	61	VSYIYIINAAAPLFLSPQIIRSPILNIRSLIRKILVSWMTPEPYPLGISMISAISTER	120
Qy	136	CSLTIIMPWYHGRPRPRYIASSVWCYVLLMALSLIRSLIEMWFCDFLPSGADSWCETSDPIT	195
Db	121	CSLVTLPWYRCRRPPTHLSAVVCYVLLMGLSLIFSLMEWRFCDFLPSGADSWCETSDPIT	180
Qy	196	IATLVAVLCYVLCGSSLYLVYRLTLCGRKMPILTRLYVYITLTVLVPFLCGLPFGIQWMLFS	255
Db	181	VWVLVPLCYVLLCASSLVAVLRLLCGSRKMPILTRLYVYITLTVLVPFLCGLPFGILGHLIY	240
Qy	256	RILHDKVLFCHVHLVSIPLSAINSSANPIIYFPVGSFRORONQMLVYLORALQDTPK	315
Db	241	RMLIANLEVYVCHVYLVCMSSLSAINSSANPIIYFPVGSFRORONQMLVYLORALQDKPS	300
Qy	316	VDEGGGMLPQETLIELSGSRL	335
Db	301	VDKGGGLPBPBSLELSGSKL	320

```

RESULT 7
US-09-254-227A-1
Sequence 1, Application US/09254227A
Patent No. 6696257
GENERAL INFORMATION:
APPLICANT: Ahmad, Sultana
APPLICANT: Banville, Denis
APPLICANT: Fortin, Yves
APPLICANT: Lembo, Paola
APPLICANT: O'Donnell, Dajan
APPLICANT: Shi-Hsiang, Shen
TITLE OF INVENTION: G Protein-Coupled Receptors from the Rat and Human
FILE REFERENCE: 81823/266117
CURRENT APPLICATION NUMBER: US/09/254,227A
CURRENT FILING DATE: 1999-03-03
NUMBER OF SEQ ID NOS: 22
SOFTWARE: Patentin version 3.0
SEQ ID NO 1
LENGTH: 337
TYPE: PRT
ORGANISM: rat
US-09-254-227A-1

```

Query Match	47.3%	Score 834.5	DB 4	Length 337
Best Local Similarity	52.9%	Pred. No. 3.3e-65		
Matches 174	Conservative 52	Mismatches 92	Indels 11	Gaps 6

```
Qy      13 FLSNDSTPIVIGTELTPINGBETPCYKQTLSFTGLTCLIVSLVALTGNAVVLMLGCRMR 72
        ||||| : ||| : | : | : ||||| ||
Db      12 FVSMDPTISSLSTESTTLNKGTHPSC-RPILTISFLVPITITLLAGNTIVLMMLGFMRM 70
```


COUNTRY: USA
ZIP: 20004
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: Patentin Release #1.0, Version #1.25
CURRENT APPLICATION DATA:
APPLICATION NUMBER: PCT/US93/08528
FILING DATE: 09-SEP-1993
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 07/943,236
FILING DATE: 10-SEP-1992
ATTORNEY/AGENT INFORMATION:
NAME: Townsend, Kevin G.
REGISTRATION NUMBER: 34,033
REFERENCE/DOCKET NUMBER: MURPHY=2 PCT
TELECOMMUNICATION INFORMATION:
TELEPHONE: 202-628-5197
TELEFAX: 202-737-3528
TELEX: 248633
INFORMATION FOR SEQ ID NO: 52:
SEQUENCE CHARACTERISTICS:
LENGTH: 282 amino acids
TYPE: amino acid
STRANDEDNESS: single
TOPOLOGY: linear
MOLECULE TYPE: peptide
PCT-US93-08528-52

Query Match 22.0%; Score 388.5; DB 5; Length 282;

Best Local Similarity 34.4%; Pred. No. 9e-27; Matches 96; Conservative 53; Mismatches 87; Indels 43; Gaps 7;

QY 52 VSLVALTGNVAVLM-----LLGCRNRNAVSIYIINTL-VAADFLFLSG 93
DB 9 ISPVGVFENGILLMFLCFYVYTHLSIADISLFC-----IFLSIDYALDYELSLG 60
QY 94 HIICSPRLININHPISKISIPWTFPPYIGLSMLSAISTERCLSIIMPTWYHCRRRYL 153
DB 61 H-----YTVITLSVTFPLFGYNTGLYLTATISVERCLSVLYPIWYRCHRPKQ 108
QY 154 SSVMCVLWALSLRSILSLEMMFCDFPSGADSVWCETSDFTITAMVFLCVLGGSLVL 213
DB 109 SALVCALLMLSLCIVTM-YVMCIDRPEBSHRDCAVILFNLISFLVPTSVSSTIL 167
QY 214 LVRIILGSRKMPRLRYVITLTVLVLGCLPFGIOMALFSRIHLDWKVLFCHVHLVSI 273
DB 168 VVKIRKWTASHSSKLYIVIMVTIIFLIFAMPWRLLYLTYER--WST-FGMLHHSIL 223
QY 274 FLSALNSSANPIIYFVGSFRORONRONTKLVLODPEVDEGGMLPQD 312
DB 224 LFSITINSSANPIIYFVGSFKKRFKESLKVLTFRARFD 262

RESULT 11

US-08-118-270-76
Sequence 76, Application US/08118270
Patent No. 5508384
GENERAL INFORMATION:
APPLICANT: Murphy, Randall B.
TITLE OF INVENTION: POLYPEPTIDES OF G-COUPLED PROTEIN
TITLE OF INVENTION: RECEPTORS, AND COMPOSITIONS AND METHODS THEREOF
NUMBER OF SEQUENCES: 348
CORRESPONDENCE ADDRESSES:
ADDRESSES: BROMDY AND NEIMARK
STREET: 419 Seventh Street, N.W., Suite 300
CITY: Washington
STATE: D.C.
COUNTRY: USA
ZIP: 20004
COMPUTER READABLE FORM:
CURRENT APPLICATION DATA:

MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: Patentin Release #1.0, Version #1.25
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/118,270
FILING DATE: 09-SEP-1993
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 07/943,236
FILING DATE: 10-SEP-1992
ATTORNEY/AGENT INFORMATION:
NAME: Townsend, Kevin G.
REGISTRATION NUMBER: 34,033
REFERENCE/DOCKET NUMBER: MURPHY=2A
TELECOMMUNICATION INFORMATION:
TELEPHONE: 202-628-5197
TELEFAX: 202-737-3528
TELEX: 248633
INFORMATION FOR SEQ ID NO: 76:
SEQUENCE CHARACTERISTICS:
LENGTH: 298 amino acids
TYPE: amino acid
STRANDEDNESS: single
TOPOLOGY: linear
MOLECULE TYPE: peptide
US-08-118-270-76

Query Match 19.5%; Score 344; DB 1; Length 298;

Best Local Similarity 31.3%; Pred. No. 8.4e-23; Matches 94; Conservative 58; Mismatches 114; Indels 34; Gaps 11;

QY 48 LTCIVSLVALTGNVAVLMMLGCRNRNAVSIYI--LNLVAADFLPFGHICSPRLINI 105
DB 8 LTLCLGLV--GNGVLMPFGESIKTPTPSIYIFLHISADDIYLFSAK---ILLNLN 61
QY 106 RHPISKI-----LSPWTFPPYIGLSMLSAISTERCLSIIMPTWYHCRRRYLSSVMCV 159
DB 62 GTFGLSPDPYVRVRSIVGLTFPAGVSLPLAISIERCVSIYFPMWYRRRPKLSAGVCA 121
QY 160 LMLALSILSLSLEMMFCDFPSGADSVWCETSDFTITAMVFL-----LCVYLCGSSSLVL 214
DB 122 LMLLFLVLYTSIHNYFC-LIGHERASGTACIMNDISLIGLILFLFCPIWVLPD---IAL 176
QY 215 VRIILGSRKMPRLRYVITLTVLVLGCLPFGIOMALFSRIHLDW--KVLFCVHLV 271
DB 177 LHVECARRRQSAKANHVLAIVSVFLVSSITGLDMFLF-----WVQIAPPEYV 230
QY 272 SIFLSALNSSANPIIYFVGSFRORONRONTKLVLODPEVDEGGMLPQD 312
DB 231 RDLCTICINSAKPIYVFIAGRDSQRLMEPLRVVFORALRDGAEPGDASSTENTVTMEM 290

RESULT 12

PCT-US93-08528-76
Sequence 76, Application PC/TUS9308528
GENERAL INFORMATION:
APPLICANT: New York University
TITLE OF INVENTION: POLYPEPTIDES OF G-COUPLED PROTEIN
TITLE OF INVENTION: RECEPTORS, AND COMPOSITIONS AND METHODS THEREOF
NUMBER OF SEQUENCES: 348
CORRESPONDENCE ADDRESSES:
ADDRESSES: BROMDY AND NEIMARK
STREET: 419 Seventh Street, N.W., Suite 300
CITY: Washington
STATE: D.C.
COUNTRY: USA
ZIP: 20004
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: Patentin Release #1.0, Version #1.25
CURRENT APPLICATION DATA:

APPLICATION NUMBER: PCT/US93/08528
FILING DATE: 09-SEP-1993
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 07/943,236
FILING DATE: 10-SEP-1992
ATTORNEY/AGENT INFORMATION:
NAME: Townsend, Kevin G.
REGISTRATION NUMBER: 34,033
REFERENCE/DOCKET NUMBER: MURPHY=2 PCT
TELECOMMUNICATION INFORMATION:
TELEPHONE: 202-628-5197
TELEFAX: 202-737-3528
TELEX: 248633
INFORMATION FOR SEQ ID NO: 76:
SEQUENCE CHARACTERISTICS:
LENGTH: 298 amino acids
TYPE: amino acid
STRANDEDNESS: single
TOPOLOGY: linear
MOLECULE TYPE: peptide
PCT-US93-08528-76

Query Match 19.5% Score 344; DB 5; Length 298;

Best Local Similarity 31.3% Pred. No. 8,4e-23; Indels 34; Gaps 11;
Matches 94; Conservative 58; Mismatches 114;

QY 48 LITCIVSLVATGNVAVLWMLGCRMRNNAVSIIYI--LNLVAADFLPLSGHIIICSPRLINI 105
DB 8 LITCIVSLVATGNVAVLWMLGCRMRNNAVSIIYI--LNLVAADFLPLSGHIIICSPRLINI 61
QY 106 RHPLISKI-----LSPVMTFPFIFGLSMLSAISTERCUSTLMPYTHCRPRYLSSVMCV 159
DB 62 GTFPGSPDYRRVRSRIQGLTFEAGVSLPLPAISIRCVSIPEPMYWRPRRLSAGVCA 121
QY 160 LLMALSLRLSILEMWFCDPLFGSADSVWCETSDFTITAMLVF-----LCVVLGSSVLVL 214
DB 122 LLMALSLRLSILEMWFCDPLFGSADSVWCETSDFTITAMLVF-----LCVVLGSSVLVL 176
QY 215 VRLIGSRKMPLT-RLVYVITLITVLVFLGLPFGIOWALFSRHLDM--KVLFGHVLV 271
DB 177 LHVCRARRRQSAKLNNVLAIVSVFLVSSLYLGIDWLF-----WVFGIAPPEPV 230
QY 272 SIFSLANSNAPITFYFVGSFRORONRMLKVLQORALQDTPVEDEGGMLPQE-TLEL 330
DB 231 RDLICINSSAKPIVYFTAGRDQSRLWEPLRVFORALRDGAEGBDAASTPTVTWEM 290

RESULT 13

US-08-981-825-6
Sequence 6, Application US/08981825
Patent No. 6040426
GENERAL INFORMATION:
APPLICANT: OGAMA, KAZUYAUKI
APPLICANT: TANAKA, KAZUYA
APPLICANT: NAGATA, KIYUHI
APPLICANT: TAKANO, SYOICHI
TITLE OF INVENTION: PROTEIN SPECIFIC TO HUMAN
TITLE OF INVENTION: TN2, GENE (B19) ENCODING THE SAME, AND
NUMBER OF SEQUENCES: 6
CORRESPONDENCE ADDRESS:
ADDRESSEE: Knobb, Martens, Olson & Bear
STREET: 620 Newport Center Drive 16th Floor
CITY: Newport Beach
STATE: CA
COUNTRY: U.S.A.
ZIP: 92660
COMPUTER READABLE FORM:
MEDIUM TYPE: Diskette
COMPUTER: IBM Compatible
OPERATING SYSTEM: DOS
SOFTWARE: FastSeq Version 1.5
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/981,825

FILING DATE: 530
CLASSIFICATION: 530
PRIOR APPLICATION DATA:
APPLICATION NUMBER:
FILING DATE:
ATTORNEY/AGENT INFORMATION:
NAME: Altman, Daniel E.
REGISTRATION NUMBER: 34,115
REFERENCE/DOCKET NUMBER: MSHIM4.001APC
TELECOMMUNICATION INFORMATION:
TELEPHONE: 714-760-0404
TELEFAX: 714-760-9502
TELEX:
INFORMATION FOR SEQ ID NO: 6:
SEQUENCE CHARACTERISTICS:
LENGTH: 395 amino acids
TYPE: amino acid
STRANDEDNESS: unknown
TOPOLOGY: unknown
MOLECULE TYPE: protein
US-08-981-825-6

Query Match 14.5% Score 255.5; DB 3; Length 395;

Best Local Similarity 25.9% Pred. No. 8e-15; Indels 75; Gaps 10;
Matches 83; Conservative 56; Mismatches 107;

QY 51 IVSLVATGNVAVLWMLGCRMRNNAVSIIYI--LNLVAADFL-----FLS-----G 93
DB 41 LSLILGIVNCGVILFVVGCRMRQVTTVTLHLASLPLFTYFLAVGHSWELG 100
QY 94 HICSPPLIRIRPISKILSPVMTFPFIFGLSMLSAISTERCUSTLMPYTHCRPRYL 153
DB 101 TTFC-----KLHSSIFPLNMFAGFLSAISIDRCIQVPRVWQNHRTVAA 147
QY 154 SSMCVLLMALSLRLSILEMWFCDPLFGSADSVWC-----E 189
DB 148 AKKVLVLMALVANTVYFVFRITISRLDRICVYNVLLNPGPRDRTCSROALA 207
QY 190 TSDPTITAMLVFLCVLGGSLVLLVNLGSRMRPLTRLVYVITLITVLVFLGLPFGI 249
DB 208 VSKFL-LAFVLPLAIIASHAAVSL-RQHRGRRP-GRFVRLVAAVAAVAFALCWGPHYV 264
QY 250 QMALFSRIH-----LDKVLFGHVLVSLFSLAL--NSSANPIFYFVGSFRORONR 299
DB 265 FSLLEARRAHANPGLRPLWRGL-----PFTSLAFNSVANPVLVYITCPDMLRKL 316
QY 300 ONLKLVLQORALQDTPVEDEGG 320
DB 317 RSLRTVLESVLVDSELGAG 337

RESULT 14

US-09-480-784-6
Sequence 6, Application US/09480784
Patent No. 6166186
GENERAL INFORMATION:
APPLICANT: OGAMA, KAZUYAUKI
APPLICANT: TANAKA, KAZUYA
APPLICANT: NAGATA, KIYUHI
APPLICANT: TAKANO, SYOICHI
TITLE OF INVENTION: PROTEIN SPECIFIC TO HUMAN
TITLE OF INVENTION: TN2, GENE (B19) ENCODING THE SAME, AND
NUMBER OF SEQUENCES: 6
CORRESPONDENCE ADDRESS:
ADDRESSEE: Knobb, Martens, Olson & Bear
STREET: 620 Newport Center Drive 16th Floor
CITY: Newport Beach
STATE: CA
COUNTRY: U.S.A.
ZIP: 92660
COMPUTER READABLE FORM:
MEDIUM TYPE: Diskette
COMPUTER: IBM Compatible

This Page Blank (uspio)

GenCore version 5.1.6
Copyright (c) 1993 - 2004 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: October 27, 2004, 08:25:18 ; Search time 39 Seconds

(without alignments)
831.412 Million cell updates/sec

Title: US-09-867-570-2

Perfect score: 1763
Sequence: 1 MESKSSWVIRLGLFSLMDSTL.....EGGWLPGQTLRLSGRLAQ 337Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 283416 seqs, 96216763 residues

Total number of hits satisfying chosen parameters: 283416

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Listing first 45 summaries

Database :

1: p1r1:*
2: p1r2:*
3: p1r3:*
4: p1r4:*

Pred. No. is the number of results predicted by chance to have a
score greater than or equal to the score of the result being printed,
and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	461	26.1	378	A39485	transforming prote
2	446.5	25.3	324	1 TVRTAS	transforming prote
3	430.5	24.4	325	1 TVHVAS	transforming prote
4	428.5	24.3	324	2 S51001	transforming prote
5	371	21.0	343	2 A35639	G protein-coupled
6	240	13.6	355	2 JQ1231	interleukin-8 rece
7	238	13.5	353	2 C42009	interleukin-8 rece
8	235.5	13.4	354	2 A23669	interleukin-8 rece
9	226.5	12.8	360	2 A53611	interleukin-8 rece
10	220	12.5	358	2 A53752	interleukin-8 rece
11	219.5	12.5	369	2 D41795	somatostatin recep
12	219.5	12.5	369	2 A45291	somatostatin recep
13	217	12.3	351	2 B42009	PMIP-related recep
14	214.5	12.2	346	2 S29248	somatostatin recep
15	214.5	12.2	388	2 JN0605	somatostatin recep
16	210	11.9	473	2 JC5835	anaphylatoxin C3a
17	209.5	11.9	363	2 I57940	somatostatin recep
18	209	11.9	356	2 S42096	interleukin-8 rece
19	206.5	11.7	384	2 A47249	brain-specific som
20	206	11.7	369	2 JC2083	somatostatin recep
21	206	11.7	380	2 S36143	kappa opioid recep
22	205.5	11.7	355	2 A55733	G protein-coupled
23	205	11.6	333	2 I65989	kappa opioid recep
24	204	11.6	380	2 A48227	kappa opioid recep
25	201.5	11.4	380	2 A55259	kappa opioid recep
26	201	11.4	352	2 A46520	N-formyl peptide r
27	201	11.4	380	2 JC2434	kappa opioid recep
28	199	11.3	369	2 B41795	somatostatin recep
29	198.5	11.3	380	2 JC2338	kappa opioid recep

30	196.5	11.1	350	1 A37963	complement C5a ana
31	196.5	11.1	371	2 JC5498	G protein-coupled
32	196.5	11.1	504	2 A41783	tachykinin recepto
33	196	11.1	384	2 JC4629	somatostatin recep
34	194.5	11.0	350	2 A39445	interleukin-8 rece
35	194	11.0	363	2 I57955	somatostatin recep
36	194	11.0	364	2 JN0763	somatostatin recep
37	193.5	11.0	375	2 JC5069	G protein-coupled
38	193	10.9	364	2 A49542	N-formyl peptide c
39	189	10.7	392	2 S65693	opioid receptor mu
40	188	10.7	359	2 I49341	MIP-1 alpha recept
41	187.5	10.6	352	1 S27357	complement C5a ana
42	187	10.6	340	2 JC7695	G protein-coupled
43	187	10.6	378	2 A55735	G protein-coupled
44	186	10.6	391	2 A41795	somatostatin recep
45	186	10.6	391	2 C41795	somatostatin recep

ALIGNMENTS

RESULT 1

A39485 transforming protein (mxg) - human

C/Species: Homo sapiens (man)

C/Date: 28-Feb-1992 #sequence_revision 17-Apr-1993 #text_change 09-Jul-2004

C/Accession: A39485

R/Komoc, C.; Weber, V.; Stimmack, J.; Bihoreau, C.; Teutsch, B.; Corvol, P.; Clauser, Mol. Endocrinol. 5, 1477-1487, 1991

A/Title: Cloning and functional characterization of a novel mas-related gene, modulating

A/Reference number: A39485; MUID:92130997; PMID:1723144

A/Accession: A39485

A/Status: preliminary

A/Molecule type: DNA

A/Residues: 1-378 <MON>

A/Cross-references: UNIPROT:P35410; GB:S78653; NID:G244209; PIDN:AM21255.1; PID:G244210

C/Superfamily: mas transforming protein

C/Keywords: G protein-coupled receptor; transmembrane protein

Query Match 26.1%; Score 461; DB 2; Length 378;

Best Local Similarity 37.9%; Pred. No. 2.7e-31;

Matches 107; Conservative 56; Mismatches 81; Indels 38; Gaps 10;

QY	51	IVSLVLTGNVAVLMLGCMRRRAVSITLNLVAADFLPSGHILCSPLRLINR----	106
DB	84	LVSLGVLNGVFWMLCCG-ATNPVNYTLHLVAADVTL-----CSAVGFLQVTLTLTY	138
QY	107	HPI-----SKLSPVTPPEYFGSLMSAISTERCLSTLMPVYHCRPRRLSYWCV	159
DB	139	HGVVFPIDPLALISF--FSFVCLCLVAISTERCVLPVWRGHRPKYSNVCT	195
QY	160	LWALSLRSILEWMPDFPSGADSVWCETSD--FITAML--VFLCVLCCSSLVLL	214
DB	196	LWGLPFCINIVKSLFLTY-----WKHVKAQVIFLKLSGFHALSLVMCVSSYTLT	247
QY	215	VRILCSRRKMLPRLVYTLTLVLPFLCGLRPGIQWALSRLHLDKVLFCHEHLVSIF	274
DB	248	IRFLCSQOQKARIVAVVOISAPWFLMLADPLSV-----APLTTDKMFTVTSYLSLF	302
QY	275	LSALNSSANPIIYFVSGFRQQRQWKLVLQALADTPREV	316
DB	303	L-INSSANPIIYFVSGSLKKRLKESLRYLQALADKREV	343

RESULT 2

TVRTAS transforming protein mas - rat

C/Species: Rattus norvegicus (Norway rat)

C/Date: 31-Dec-1989 #sequence_revision 31-Dec-1989 #text_change 09-Jul-2004

C/Accession: A31816

R/Young, D.; O'Neill, K.; Jessell, T.; Wigler, M. Proc. Natl. Acad. Sci. U.S.A. 85, 5339-5342, 1988

A/Title: Characterization of the rat mas oncogene and its high-level expression in the h

Db 272 SSANPFIYFVSGSSKKRFRSLKVLVLTAFKDEMQRROGNG 315

RESULT 5

A:35639 G protein-coupled receptor RTA - rat

C:Species: Rattus norvegicus (Norway rat)

C>Date: 22-Jan-1993 #sequence_revision 22-Jan-1993 #text_change 09-Jul-2004

C/Accession: A35639

R/Ross, P.C.; Figliet, R.A.; Corjay, M.H.; Barber, C.M.; Adam, N.; Marcus, D.R.; Lynch, K.

Proc. Natl. Acad. Sci. U.S.A. 87, 3052-3056, 1990

A>Title: RTA, a candidate G protein-coupled receptor: cloning, sequencing, and tissue di

A/Reference number: A35639; MUID:90222168; PMID:2109324

A/Accession: A35639

A:Molecule type: mRNA

A/Residues: 1-343 <ROS>

A/Cross-references: UNIPROT:P23749; GB:M35297; NID:G206809; PID:AAA42087.1; PID:G206810

C/Superfamily: mas transforming protein

C/Keywords: G protein-coupled receptor; glycoprotein; membrane protein

F/Binding site: carbohydrate (Aen) (covalent) #status predicted

Query Match 21.0%; Score 371; DB 2; Length 343;

Best Local Similarity 33.1%; Pred. No. 9e-24;

Matches 99; Conservative 58; Mismatches 110; Indels 32; Gaps 11;

QY 48 LFCIVSLVNLGNVAVMLGCGRRRNVASITLNLVAADFLPSGHIIICSPRLNIR- 106

Db 53 LCLCLGLV--GGLVLMFPGFSTKRTPSIFYLHLASADGIYFSAV--IALLMNGT 106

QY 107 -----HPISKILSPVWTPPYFIFGLSMLSAISTERCLSIPIWYHCRPRYLSSVM 157

Db 107 FLGSPDYVRVRVRIIG-LCTF--FAGVSLIPALISIRCVSVIFPMYWRRRPRLSAGV 163

QY 158 CVLLMALSLRLSILEMWFCDPLFSGADSWCETSDFTITAMLVF--LCVLCSSGLVLLV 215

Db 164 CALLMLSLFVLTSTIHNFCFLGHEASGTACLMMND-ISTGILFLFLCPLMVLPCLALIL 222

QY 216 RIICGSRKMLT-RLVYTTILTVLVLLCGLPFGIOMALSRHLDK--KVLPGHVLVS 272

Db 223 HVCCRARRRORSAKLNHVLAIVSVFLVSSITVIGIDIMFLF--VVFQIPADPPRYVT 276

QY 273 IFLSALNSANPFIYFVSGFRORONRQNLKLVQRLADTPVEDSGGMLPQE-TLEL 330

Db 277 DLICINSSAKPIYFLAGRDQGRMLERLVRFQRLRQAGARGDAASTPNVTMEM 335

RESULT 6

interleukin-8 receptor - rabbit

C:Species: Oryctolagus cuniculus (domestic rabbit)

C>Date: 31-Mar-1992 #sequence_revision 31-Mar-1992 #text_change 09-Jul-2004

C/Accession: JQ1231; A46483

R/Beckmann, M.P.; Munger, W.E.; Kozlosky, C.; Vandenberg, T.; Price, V.; Lyman, S.; Gerard

Biochem. Biophys. Res. Commun. 179, 784-789, 1991

A>Title: Molecular characterization of the interleukin-8 receptor.

A/Reference number: JQ1231; MUID:91378994; PMID:1898400

A/Accession: JQ1231

A:Molecule type: DNA

A/Residues: 1-355 <BEC>

A/Cross-references: UNIPROT:P21109; GB:M74240; NID:G165438; PID:AAA1375.1; PID:G165439

R/Ree, J.; Khang, W.J.; Rice, G.C.; Wood, W.I.

J. Immunol. 148, 1261-1264, 1992

A>Title: Characterization of complementary DNA clones encoding the rabbit IL-8 receptor.

A/Reference number: A46483; MUID:92148149; PMID:11737938

A/Accession: A46483

A/Status: preliminary

A:Molecule type: mRNA

A/Residues: 1-355 <LEB>

A/Cross-references: GB:M82873; NID:G165440; PID:AAA1376.1; PID:G165441

A/Experimental source: neutrophil

A/Note: sequence extracted from NCBI backbone (NCBIN:81526, NCBI:P.81530)

C/Superfamily: vertebrate rhodopsin

C/Keywords: G protein-coupled receptor; transmembrane protein

Query Match 13.6%; Score 240; DB 2; Length 355;

Best Local Similarity 27.4%; Pred. No. 9.4e-13;

Matches 92; Conservative 57; Mismatches 109; Indels 78; Gaps 17;

QY 25 TELTPINGREBTPC--YKQTLSS--FTGLTICVSLVLTGNA-VLWMLGCRMRNNAVSI 78

Db 23 TGMPEPVE-KOYSPCLVVTQTLNKVAVVIAVLFSLSLNSLVMLVILVSRNSRSTVD 81

QY 79 YILNLVAADFLPSGHIIICSPRLIN-----IRHPSKILSPVWTPPYFIFGLSMLSAIS 132

Db 82 YILNLVAADFLP-----ALTPRINAVSKEKMWIRCTPLCKVSLVKEVNFYSGILLACIS 137

QY 133 TERCLSLMPDIWYHCR-----PRYSVWCVLLMALSLRLSILEMWFCDPLFSGADSWVC 188

Db 138 VDRYIATV-----HATTLTQKRLHVFLICGLWALSILSPFLRQVFPNNSSPVC 192

QY 189 -ETSDFITIAMLVFLCVV--LCGSSLVLTAVRILC-GSRKMPRLVY-----TILLT 236

Db 193 YEDLGHNTAKRMVRLRILPHTFGRIPLVLMFCYGTLLRTLFGAHNGQKRAMRVIFAV 252

QY 237 VLVFLCGLPFGIOMALFSRIHLDKVLFC--HNLV-----SIFLS 276

Db 253 VLIPLCLMLPYNL-----VLADDTLMRTVHIGETCQRNDIDRALDATEILG 299

QY 277 ALNSANPFIYFVSGFRORONRQNLKLVQRLADTPVEDSGGMLPQE-TLEL 310

Db 300 FLHSCINPIYAFIG-----QFRNGFLKMLAARGL 330

RESULT 7

FMPL-related receptor 2 - human

N/Alternate names: FMPL-related receptor 1; probable chemotactic receptor FPRH2

C:Species: Homo sapiens (man)

C>Date: 30-Sep-1993 #sequence_revision 30-Sep-1993 #text_change 09-Jul-2004

C/Accession: C42009

R/Bao, L.; Gerard, N.P.; Eddy Jr., R.L.; Shows, T.B.; Gerard, C.

Genomics 13, 437-440, 1992

A>Title: Mapping of genes for the human C5a receptor (C5AR), human FMPL receptor (FPR),

A/Reference number: A42009; MUID:92307681; PMID:1612600

A/Accession: C42009

A/Status: nucleic acid sequence not shown

A:Molecule type: DNA

A/Residues: 1-353 <BAO>

A/Cross-references: UNIPROT:P25089; GB:M76673; NID:G182668; PID:G182669

C/Comment: This fMet-Leu-Phe receptor homolog, whose ligand is not yet known, appears nc

C/Genetics: A:Gene: GDB:FPRL2

A/Cross-references: GDB:128855; OMIM:136539

A/Map position: 19q13.3-19q13.4

A/Intons: #status absent

A/Superfamily: vertebrate rhodopsin

C/Keywords: Chemotaxis; G protein-coupled receptor; glycoprotein; transmembrane protein

Query Match 13.5%; Score 238; DB 2; Length 353;

Best Local Similarity 23.8%; Pred. No. 1.4e-12;

Matches 81; Conservative 76; Mismatches 112; Indels 72; Gaps 15;

QY 29 PINGREB-----TPCYQTLSTFTGLTICVSLVLTGNAVLMGCRMRNNAVSIYI 80

Db 8 PLMBTEFVLBPBAGHTVLMIFSLVHGVTFVGL--GGLVIVWAGFRTTRVNTICY 64

QY 81 LNLVAADFLPSGHIIICSPRLINI---RHP-----ISKILSPVWTPPYFIFGLSMLSAIS 132

Db 65 LNLVAADFLPSGA---ILPRMVSVAAREKMPFASCKVHVWIDINLFPVSVLITIIA 120

QY 133 TERCLSLMPDIWYHCRPRYLSSVMCVLLMALSLRLSILEMWF-----CDPLF 180

Db 121 LDRCTCVLHPAMQNHRTMSLAKRVMTGLMIFITVLTPNFIFTTISTTNGDYCLFNF 180

QY 181 SGADSWCET-----SDFITIAMLVFLCVVLCGSSLVLTAVRILC-----G 220

Db 181 ----AFMGDTAVERLNVETMAKVELLHFIIGTVPMSIITVCYGIINAKIRHNMKS 236
 QY 221 SRKMPLERLVYTTILTVLVLGCPF---GIQWLFSPR---IHLDMKVLFCVHLVSI 274
 Db 237 SR--PL-RVFAAV---VASFTICMPFYEIGILMAWMLKEMILNGKIKIILVILNPTS-S 289
 QY 275 LSAINSSANPIITYFVSGFRORONRKLVLVQRLADPTPE 315
 Db 290 LAFNNSCLNPIILYVWGNFQRLRLSLPTSLERLALTEVPD 330

RESULT 8

A23669
 Interleukin-8 receptor, high affinity - rabbit

N:Alternate names: FMLP receptor

C:Species: Oryctolagus cuniculus (domestic rabbit)

C>Date: 22-Jan-1993 #sequence_revision 22-Jan-1993 #text_change 09-Jul-2004

C/Accession: A23669

R/Thomas, K.M.; Pyun, H.Y.; Navarro, J.

J. Biol. Chem. 265, 20061-20064, 1990

A>Title: Molecular cloning of the Fmet-leu-Phe receptor from neutrophils.

A/Reference number: A23669; MUID:91056034; PMID:1700779

A/Accession: A23669

A/Molecule type: mRNA

A/Residues: 1-354 <THO>

A/Cross-references: UNIPROT:P21109; GB:M58021; GB:J05705; NID:g165442; PIDN:AAA1377.1;

C:Superfamily: vertebrate rhodopsin

C/Keywords: G protein-coupled receptor; glycoprotein; membrane protein; neutrophil

Query Match 13.4%; Score 235.5; DB 2; Length 354;
 Best Local Similarity 28.9%; Pred. No. 2.2e-12;
 Matches 94; Conservative 53; Mismatches 121; Indels 57; Gaps 16;

QY 25 TELPINGERETPC--YKOTLS---FTGLTCVSLVALTGN--VVLMLGCMRRNAVSI 78
 Db 23 TGMPEV--KDYSPCLVVTQTLNKYVVVYVALVFLSLGSLVMLVILYSRNSVTDV 81
 QY 79 YILNLVADFLPLSGHILCSPL--RLINIRHPIKILSPVWTFPFYIGLSMLSAISTERC 136
 Db 82 YILNLVADFLPLSGHILCSPL--RLINIRHPIKILSPVWTFPFYIGLSMLSAISTERC 140
 QY 137 LSLIMPIWYHCRPRYLSVNCVLLMALSILRSILEMFCDFLFGADSVWC--ETSDPTT 195
 Db 141 LAIVQST--RLTLQKHILVKKFICLGIWALSILSLPFLFRQVSPNNSSPVCYEDIGHNT 199
 QY 196 IAWLVPLCVL-----GSSSLVILVRLIC--GSRKMPLTRLV-----TILTVAYE 240
 Db 200 AKM---CMVLRILPHRTGFLPLVLMFLCYGFTLRLEFQAMGQKRAMRVIPAVLVIF 255
 QY 241 LILGLPFG-----IQWALFSRIHLDKVLFCVHLVSIPLSALNSSANPIIY 287
 Db 256 LILGLPFG-----IQWALFSRIHLDKVLFCVHLVSIPLSALNSSANPIIY 309
 QY 288 FVSGFRORONR--LKLVLQRL 310
 Db 310 AFIG-----QMFNGFLKMLAARGL 329

RESULT 9

A53611

Interleukin-8 receptor type B - human

C/Species: Homo sapiens (man)

C>Date: 07-Oct-1994 #sequence_revision 12-Apr-1996 #text_change 09-Jul-2004

C/Accession: I37898; I38712; A53611; A39446

R/Hubb, S.K.; Shetty, A.; Tiffany, H.L.; Murphy, P.M.

J. Biol. Chem. 265, 26381-26389, 1994

A>Title: Comparison of the genomic organization and promoter function for human interleukin-8 receptor type B and human interleukin-8 receptor type A.

A/Accession: I37898

A/Status: preliminary

A/Molecule type: DNA

A/Residues: 1-360 <RBS>
 A/Cross-references: UNIPROT:P25025; EMBL:U11869; NID:g511801; PIDN:AA60656.1; PID:g5118

A/Accession: I38712
 A/Status: preliminary
 A/Molecule type: mRNA
 A/Residues: 1-15 <RBS>
 A/Cross-references: EMBL:U11872; NID:g511808; PIDN:AAA4380.1; PID:g511809; EMBL:U11873;
 I1876; NID:g511816; PID:g511817; EMBL:U11877; NID:g511818; PID:g511819; EMBL:U11878; NID
 J. Biol. Chem. 269, 11065-11072, 1994
 A>Title: Structure, genomic organization, and expression of the human interleukin-8 recep
 A/Reference number: A53611; MUID:94209273; PMID:7512557

A/Accession: A53611

A/Status: preliminary

A/Molecule type: DNA

A/Residues: 6-360 <SPR>

A/Cross-references: GB:M99412; GB:I19593

R/Murphy, P.M.; Tiffany, H.L.

Science 253, 1280-1283, 1991

A>Title: Cloning of complementary DNA encoding a functional human interleukin-8 receptor

A/Reference number: A39446

A/Accession: A39446

A/Status: preliminary; nucleic acid sequence not shown

A/Molecule type: mRNA

A/Residues: 6-360 <MUR>

A/Cross-references: GB:M73969

C/Comment: This receptor, unlike IL8RA, binds several peptides besides interleukin-8, inc

C:Genetics:

A/Map position: 2q35-2q35

C:Superfamily: vertebrate rhodopsin

C/Keywords: G protein-coupled receptor; transmembrane protein

Query Match 12.8%; Score 226.5; DB 2; Length 360;
 Best Local Similarity 26.3%; Pred. No. 1.3e-11;
 Matches 82; Conservative 52; Mismatches 103; Indels 75; Gaps 13;

QY 34 EETPCYKOTLSFTG---LTCVSLVALTGN--VVLMLGCMRRNAVSIYILNLVAD 87
 Db 35 DAAPCEPSLEINKYFVYIYVALVFLSLGSLVMLVILYSRNSVTDVYILNLVAD 94
 QY 88 PLP-----LSGHILCSPLRLINIRHPIKILSPVWTFPFYIGLSMLSAIS 132
 Db 95 LILFALTPLIWAASKYNGMIFGFLC-----KVSLSLKEVNFYSGILLACTS 141
 QY 133 TERCLSLPIWYHCR---PRYSVNCVLLMALSILRSILEMFCDFLFGADSVWC 188
 Db 142 VDRYLAIV---HATRLTLQKHILVKKFICLSINGSLILALPVLLFRITYSSNVSPAC 196
 QY 189 -ETSDPTTAWLVPLCVL--LCGSSLVILVRLIC--GSRKMPLTRLV-----TILTV 236
 Db 197 YEDMGNTNANWMLRLILPQSGFVPLPILMFCYGFTRLFLFKAMGQKRAMRVIPAV 256
 QY 237 VLVLGLPFG-----IQWALFSRIHLDKVLFCVHLVSIPLSALNSSAN 283
 Db 257 VLIFLGLPFG-----IQWALFSRIHLDKVLFCVHLVSIPLSALNSSAN 310
 QY 284 PIIFYEVG-SFR 294
 Db 311 PLIYAFIQGKFR 322

RESULT 10

A53752

Interleukin-8 receptor (clone 5B1a) - rabbit

C/Species: Oryctolagus cuniculus (domestic rabbit)

C>Date: 06-Jan-1995 #sequence_revision 06-Jan-1995 #text_change 09-Jul-2004

C/Accession: A53752

R/Pardo, G.N.; Thomas, K.M.; Suzuki, H.; LaRosa, G.J.; Wilkinson, N.; Folco, E.; Navarro

J. Biol. Chem. 269, 12391-12394, 1994

A>Title: Molecular characterization of a novel rabbit interleukin-8 receptor isoform.

A/Reference number: A53752; MUID:94230294; PMID:8175642

A/Accession: A53752

A/Status: preliminary

A:Molecule type: mRNA
A:Residues: 1-358 <PRA>
A/Cross-references: UNIPROT:P35344; GB:I24445; NID:g437661; PID:AAA31378.1; PID:g437662
C/Superfamily: vertebrate rhodopsin
C/Keywords: G protein-coupled receptor; transmembrane protein

Query Match 12.5%; Score 220; DB 2; Length 358;
Best Local Similarity 26.5%; Pred. No. 4.5e-11;
Matches 84; Conservative 53; Mismatches 112; Indels 68; Gaps 14;

QY 25 TELPPIINGREPTCYKQTLSTFTGLTCIVS-----LVALTGNA-VVIMLCCGRMRNAVS 78
DB 25 TDLPP-TLDSAPCRSRSLSTNSVYVILYIVLFLSLGNSLVMLVILYSRSTCVTDV 83
QY 79 YIINLVADPLFLSGHICGPRILINIRH-----PISKLSPMTTPPYIGLSMLSAIS 132
DB 84 YIINLVADPLFLSGHICGPRILINIRH-----PISKLSPMTTPPYIGLSMLSAIS 139
QY 133 TERCLSTLMPDWCR-----PRYSVMCVLMLALSILRSILEMFCDFLFGSGADSVWC 188
DB 140 VDRYLAIV-----HATTMIOKRLVYFICLSMVGSLIISPLIRNAIFPNNSPVC 194
QY 189 -ETSDPFTIAMVFLCVV--LCGSSVLVLRILC-----GSRKMPPLRLYVT 232
DB 195 YEDMGNSSTAKRMYRLPQTFPIPLVLMFCYVFTLRLTLPQAHMGQRH---RAMRV 250
QY 233 ILTTLVFLVLCGLPFGICQALFRIHLDKVLFCYHNL-----VSIFLSAL 278
DB 251 IFVAVVLFLLCWLPLYNL-----VLLTDLTMRTHVIOETCERRNDIDALDAITELIGFL 303
QY 279 NSSANPIYFVGV-SFR 294
DB 304 HSCLNPIIYAFIGQKFR 320

RESULT 11

41795
somatostatin receptor 2 - mouse
C/Species: Mus musculus (house mouse)
C/Date: 30-Sep-1993 #sequence_revision 30-Sep-1993 #text_change 09-Jul-2004
C/Accession: D41795; I56236
R/Yamada, Y.; Post, S.R.; Wang, K.; Tager, H.S.; Bell, G.I.; Selino, S.
Proc. Natl. Acad. Sci. U.S.A. 89, 251-255, 1992
A/Title: Cloning and functional characterization of a family of human and mouse somatostatin receptors
A/Reference number: A41795; MUID:92108031; PMID:1346068
A/Accession: D41795
A/Status: nucleic acid sequence not shown
A:Molecule type: DNA
A:Residues: 1-369 <YAM>
A/Cross-references: UNIPROT:P30875; GB:M61832; NID:g201060; PID:AAA58256.1; PID:g201061
R/Elloic, D.B.; Metwalli, A.; Blum, A.M.; Sandor, M.; Lynch, R.; Weinstein, J.V.
J. Immunol. 153, 1180-1186, 1994
A/Title: T lymphocytes isolated from the hepatic granulomas of schistosoma-infected mice
A/Reference number: I56236; MUID:94300079; PMID:7913111
A/Accession: I56236
A/Status: preliminary; translated from GB/EMBL/DBJ
A:Molecule type: mRNA
A:Residues: 99-309 <RES>
A/Cross-references: GB:G71756; NID:g560631
C/Superfamily: vertebrate rhodopsin
C/Keywords: G protein-coupled receptor; hormone receptor; transmembrane protein

Query Match 12.5%; Score 219.5; DB 2; Length 369;
Best Local Similarity 28.3%; Pred. No. 5.1e-11;
Matches 95; Conservative 55; Mismatches 131; Indels 55; Gaps 18;

QY 23 LGTELPINGREPT-PCYKQTLG--FTGLTCIVSLVLTGNAVLM-LLGCRMRNAVS 78
DB 21 LNSGLSGNSNGTEPYDNTSNVLFYFVVCVGLCGNTLYIVILRYAKKKTITNI 80
QY 79 YIINLVADPLFLSGHICGPRILINIRHPIK--ILSPVMT---FPYFGLSMLSAISTE 134
DB 81 YIINLVADPLFLSGHICGPRILINIRHPIK--ILSPVMT---FPYFGLSMLSAISTE 139

QY 135 RCLSIILMPI-----WYHCRPRYLSVWCYLLMALSLRSILEMWF-----CD 177
DB 140 RYLAIVHPIRSACK---RRR-TRKMINAVWCVSL-VLPIIMYAGLSNMGGRSCT 194
QY 178 FLFGSGADSVWCSTSDFTIAMVFLCV---VLCSSVLVLR-----ILGSRMPPLRL 229
DB 195 INMPGEGGAM--YTGFIYAFILGFLVPLTILICLYFIIRKVSIGIRVSSGRKRSK 252
QY 230 YVT--ILITVVLVLCGLP-----GIQNALFSRIHLDKVLFCYHNLVSLFSLN 281
DB 253 KTRMVSIVAVVIFPCWLPYIFNVSSVSVAISPTPAL--KGMDFV---VILTYANSC 306
QY 282 ANPIYFVGSFRORONRKLVLORALODTPEVD 317
DB 307 ANPIYAFVLSDNFKKSFQNVLCVKSGETDGERSD 342

RESULT 12

A45291
somatostatin receptor, somatostatin release-inhibiting factor receptor, SRIF receptor - 1
C/Species: Rattus norvegicus (Norway rat)
C/Date: 25-Mar-1993 #sequence_revision 18-Nov-1994 #text_change 09-Jul-2004
C/Accession: A45291
R/Kluxen, F.W.; Bruns, C.; Lubbert, H.
Proc. Natl. Acad. Sci. U.S.A. 89, 4618-4622, 1992
A/Title: Expression cloning of a rat brain somatostatin receptor cDNA.
A/Reference number: A45291; MUID:92262491; PMID:1374909
A/Accession: A45291
A/Status: preliminary
A:Molecule type: mRNA
A:Residues: 1-369 <KLUX>
A/Cross-references: UNIPROT:P30680; GB:M93273; NID:g207026; PID:AAA42165.1; PID:g207027
A/Note: Sequence extracted from NCBI backbone (NCBI:102315, NCBI:102316)
C/Superfamily: vertebrate rhodopsin
C/Keywords: G protein-coupled receptor; transmembrane protein

Query Match 12.5%; Score 219.5; DB 2; Length 369;
Best Local Similarity 28.3%; Pred. No. 5.1e-11;
Matches 95; Conservative 55; Mismatches 131; Indels 55; Gaps 18;

QY 23 LGTELPINGREPT-PCYKQTLG--FTGLTCIVSLVLTGNAVLM-LLGCRMRNAVS 78
DB 21 LNSGLSGNSNGTEPYDNTSNVLFYFVVCVGLCGNTLYIVILRYAKKKTITNI 80
QY 79 YIINLVADPLFLSGHICGPRILINIRHPIK--ILSPVMT---FPYFGLSMLSAISTE 134
DB 81 YIINLVADPLFLSGHICGPRILINIRHPIK--ILSPVMT---FPYFGLSMLSAISTE 139
QY 135 RCLSIILMPI-----WYHCRPRYLSVWCYLLMALSLRSILEMWF-----CD 177
DB 140 RYLAIVHPIRSACK---RRR-TRKMINAVWCVSL-VLPIIMYAGLSNMGGRSCT 194
QY 178 FLFGSGADSVWCSTSDFTIAMVFLCV---VLCSSVLVLR-----ILGSRMPPLRL 229
DB 195 INMPGEGGAM--YTGFIYAFILGFLVPLTILICLYFIIRKVSIGIRVSSGRKRSK 252
QY 230 YVT--ILITVVLVLCGLP-----GIQNALFSRIHLDKVLFCYHNLVSLFSLN 281
DB 253 KTRMVSIVAVVIFPCWLPYIFNVSSVSVAISPTPAL--KGMDFV---VILTYANSC 306
QY 282 ANPIYFVGSFRORONRKLVLORALODTPEVD 317
DB 307 ANPIYAFVLSDNFKKSFQNVLCVKSGETDGERSD 342

RESULT 13

B42009
FMPL-related receptor 1 - human
N/Alternate names: FMPL receptor homolog FPR2; formyl peptide receptor like-1; probable
C/Species: Homo sapiens (man)
C/Date: 30-Sep-1993 #sequence_revision 14-Jul-1995 #text_change 09-Jul-2004
C/Accession: B42009; J01521; A42492; I54751; S21581

C:Species: Homo sapiens (man)
C:Date: 30-Sep-1993 #sequence revision 30-Sep-1993 #text_change 09-Jul-2004
C:Accession: UN0605; JN0762; A47457
R:Yu, Y.; Song, J.; Bruno, J.F.; Berelowitz, M.
Biochem. Biophys. Res. Commun. 193, 648-652, 1993
A:title: Molecular cloning and sequencing of a human somatostatin receptor, hSSTR4.
A:Reference number: JN0605; MUID:9320656; PMID:8512564
A:Accession: JN0605
A:Molecule type: DNA
A:Residues: 1-388 <XUY>
A:Cross-references: UNIPROT:P31391; GB:U4856; NID:G222499; PIDN:AAA36623.1; PID:G292500
R:Yamada, Y.; Kagimoto, S.; Kubota, A.; Yasuda, K.; Masuda, K.; Someya, Y.; Ihara, Y.; I
Biochem. Biophys. Res. Commun. 195, 844-852, 1993
A:title: Cloning, functional expression and pharmacological characterization of a fourth
A:Reference number: JN0762; MUID:93384611; PMID:8373420
A:Accession: JN0762
A:Molecule type: DNA
A:Residues: 1-388 <YAM>
A:Cross-references: GB:D16826; NID:9693907; PIDN:BA04106.1; PID:9693908
R:Robner, L.; Raulf, F.; Brune, C.; Buetner, R.; Hofstaedter, F.; Schulte, R.
Proc. Natl. Acad. Sci. U.S.A. 90, 4196-4200, 1993
A:title: Cloning and characterization of a fourth human somatostatin receptor.
A:Reference number: A47457; MUID:93248256; PMID:8483934
A:Accession: A47457
A:Molecule type: DNA
A:Residues: 1-82, 'T', 84-364, 'K', 366-388 <ROH>
A:Cross-references: GB:U07833; NID:9307429; PIDN:AAA60565.1; PID:9307430
A:Note: Sequence extracted from NCBI database (NCBIN:130856, NCBIPI:130858)
C:Comment: This protein mediates the diverse actions of the tetradecapide somatostatin.
C:Genetics:
A:Gene: GDB:SSTR4
A:Cross-references: GDB:202662; OMIM:182454
A:Map position: 20p11.2-20p11.2
A:introns: #status absent
C:Superfamily: vertebrate rhodopsin
C:Keywords: G protein-coupled receptor; glycoprotein; hormone receptor; lipoprotein; pho
F:47-73/Domain: transmembrane #status predicted <TM1>
F:84-109/Domain: transmembrane #status predicted <TM2>
F:121-142/Domain: transmembrane #status predicted <TM3>
F:162-184/Domain: transmembrane #status predicted <TM4>
F:208-238/Domain: transmembrane #status predicted <TM5>
F:257-284/Domain: transmembrane #status predicted <TM6>
F:291-314/Domain: transmembrane #status predicted <TM7>
F:24/Binding site: carbohydrate (asn) (covalent) #status predicted
F:119-198/Disulfide bonds: #status predicted
F:161,253/Binding site: phosphate (Ser) (covalent) (by cAMP-dependent kinase) #status pr
F:327/Binding site: palmitate (Cys) (covalent) #status predicted

Query Match 12.2% Score 214.5; DB 2; Length 386;
Best Local Similarity 25.9% Pred. No. 1.4e-10;
Matches 95; Conservative 61; Mismatches 152; Indels 59; Gaps 14;

QY 14 LSNDSITPVLGTE-----LTPNGREET--PCYKQTLSPFGTLCT---VSLV 55
DB 1 MSAPSTLPFGEGSLGTAWPSAANASSAPAEAEAVAGPGDARAAGWVAIQCIYALVCLV 60
QY 56 ALTGNAVVLW-LIGCRNRNNAVSIYIINLVADFLF-LSGHIICSPRLINIRH-PISKI 112
DB 61 GLVGNALVIRVILRYAKMKTATNVIYINLVADFLFVLVSPFVASSAAL--RHMFGSV 117
QY 113 LS-----PVMTFFPIGSMLSAISTERCLSIWPIWTHCRPRYLSVSMCVLMAISLR 168
DB 118 LCPRAVLSVDGLNMTSTVFCLTVLSVDKYAVVHPLRAATYRRSVAKLINIGVWLASLV 177
QY 169 SLIEMFCDPL-FSGADSVNCE-----TSDFTIAMLV-----FLCVVLCT-----GS 209
DB 178 TLPIAIFADTRPARGGAVACNQPWPAWSAVFVVVTFLLGLPVLAIIGLTVLIVGK 237
QY 210 SLVLLVRIILGSRKMPILRLVYITLLTVLVFLGLGFPGLQWALFSRIHIDMKVLFCHV 269
DB 238 MRAVVALRAGWQORRRSEKKTIRLVLMVVVFLVCMPEFYV---VQLINLVVTSIDATVN 293
QY 270 LVGIFLSALNNSANPIIYFVGSFRQNRQNLKLVQRLALQDTPBVDEGGMLPQETLE 329

DB 294 HVELIISVANSKANPIILYGLS-----DNFRRSFORVLCRCCLLBGAGAREEPLD 345
QY 330 LSGSRLE 336
DB 346 YVATALK 352

Search completed: October 27, 2004, 08:35:03
Job time : 40 secs

This Page Blank (uspto)

GenCore version 5.1.6
Copyright (c) 1993 - 2004 Compugen Ltd.

OM protein - protein search, using sw model

Run on: October 27, 2004, 08:24:12 ; Search time 190 Seconds
(without alignments)
1020.533 Million cell updates/sec

Title: US-09-867-570-2
Perfect score: 1763
Sequence: 1 MESKSSWVIRLGFSLMSDSTI.....EGGWLPGQETLRLSGRLBQ 337

Scoring table: ELOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 1825181 seqs, 575374646 residues

Total number of hits satisfying chosen parameters: 1825181

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : UniProt 02:.*
1: uniprot_sprot:.*
2: uniprot_trembl:.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length DB	ID	Description
1	1688	95.7	322	1 MRG3_HUMAN	Q961B0 homo sapien
2	1663	94.3	322	1 AAH67292	AAH67292 homo sapi
3	1606	91.1	322	1 SNS2_HUMAN	Q8CDE0 homo sapien
4	1411	80.0	322	1 SNS3_HUMAN	Q8CDE0 homo sapien
5	1383	78.4	322	1 MRG1_HUMAN	Q961B2 homo sapien
6	1382	78.4	322	1 SNS5_HUMAN	Q8CDE0 homo sapien
7	1367	77.5	322	1 MRG4_HUMAN	Q961B9 homo sapien
8	989.5	56.1	330	1 MRG2_HUMAN	Q961B1 homo sapien
9	989.5	56.1	330	2 AAH63450	AAH63450 homo sapi
10	823.5	46.7	323	2 Q7TN42	Q7TN42 rattus norv
11	820.5	46.5	323	1 SNS1_RAT	Q8C491 rattus norv
12	797	45.2	322	2 Q8C1P3	Q8C1P3 mus musculu
13	782	44.4	304	1 MRGA_RAT	Q7TN49 rattus norv
14	778	44.1	331	2 Q91VE7	Q91VE7 rattus norv
15	763	43.3	338	2 Q8CDE4	Q8CDE4 mus musculu
16	763	43.3	338	2 Q91ZC2	Q91ZC2 mus musculu
17	761	43.2	304	1 MGAI_MOUSE	Q91ZC5 mus musculu
18	754	42.8	294	2 Q7TN48	Q7TN48 rattus norv
19	739.5	41.9	301	1 MGAG_MOUSE	Q91ZC6 mus musculu
20	729	41.3	302	1 MGAT_MOUSE	Q91ZC3 mus musculu
21	722.5	41.0	305	1 MGAT_MOUSE	Q91ZC3 mus musculu
22	713	40.4	321	2 Q91ZC0	Q91ZC0 mus musculu
23	703	39.9	338	2 Q91ZC3	Q91ZC3 mus musculu
24	691.5	39.2	313	1 MGAA_MOUSE	Q91ZC2 mus musculu
25	682.5	38.7	331	2 AAH64040	AAH64040 mus muscu
26	681.5	38.7	333	2 Q7TN45	Q7TN45 rattus norv
27	680.5	38.6	305	1 MGAT_MOUSE	Q91ZC4 mus musculu
28	680	38.6	304	1 MGAS_MOUSE	Q91ZC5 mus musculu
29	674	38.2	302	2 Q91ZB9	Q91ZB9 mus musculu
30	669.5	38.0	305	1 MGAB_MOUSE	Q91ZC6 mus musculu
31	638	36.2	314	2 Q7TN47	Q7TN47 rattus norv

32	626.5	35.5	353	2 Q7TN44	Q7TN44 rattus norv
33	617.5	35.0	312	2 Q91ZC1	Q91ZC1 mus musculu
34	596.5	33.8	328	2 Q7TN50	Q7TN50 rattus norv
35	578.5	32.8	245	2 Q7TN43	Q7TN43 rattus norv
36	576.5	32.7	247	2 Q7TN46	Q7TN46 rattus norv
37	568.5	32.2	330	2 Q7TN51	Q7TN51 mus musculu
38	541.5	30.7	320	1 MRGD_MACFA	Q617B6 macaca fasc
39	541.5	30.7	320	1 MRGD_MACFA	Q617B6 macaca fasc
40	506	28.7	321	1 MRGD_HUMAN	Q8CDE7 homo sapien
41	506	28.7	321	2 AAR05120	AAR05120 homo sapi
42	506	28.7	321	2 BAD20638	BAD20638 homo sapi
43	503	28.5	310	2 Q71119	Q71119 mus musculu
44	503	28.5	310	2 Q91ZB7	Q91ZB7 mus musculu
45	503	28.5	310	2 CAC86257	CAC86257 mus muscu

ALIGNMENTS

RESULT 1
MRG3_HUMAN STANDARD; PRT; 322 AA.
ID MRG3_HUMAN Q961B0; Q8TDE1;
AC Q961B0; Q8TDE1;
DT 01-OCT-2004 (Rel. 45, Created)
DT 01-OCT-2004 (Rel. 45, Last sequence update)
DT 01-OCT-2004 (Rel. 45, Last annotation update)
DE Mas-related G-protein coupled receptor member X3 (Sensory neuron-specific G-protein coupled receptor 1).
GN Name=MRG3; Synonyms=SNR1;
OS Homo sapiens (Human).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.
OX NCBI_TaxID=9606;
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE=21435808; Pubmed=11551509;
RA Dong X., Han S.-K., Zylka M.J., Simon M.I., Anderson D.J.;
RT "A diverse family of GPCRs expressed in specific subsets of nociceptive sensory neurons."
RL Cell 106:619-632(2001).
RN [2]
RP SEQUENCE FROM N.A., VARIANT ASN-169, AND TISSUE SPECIFICITY.
RX MEDLINE=21853733; Pubmed=11850634; DOI=10.1038/nrn815;
RA Lembo P.M.C., Grazzini E., Groblewski T., O'Donnell D., Roy M.-O., Zhang J., Hoffert C., Cao J., Schmidt R., Pelletier M., Labarre M., Goesslin M., Fortin Y., Benaville D., Shen S., Stroem P., Payza K., Dray A., Walker P., Ahmad S.;
RT "Proenkephalin A gene products activate a new family of sensory neuron-specific GPCRs."
RL Nat. Neurosci. 5:201-209(2002).
RN [3]
RP SEQUENCE FROM N.A.
RX TISSUE=Testis;
RC MEDLINE=22388257; Pubmed=12477932; DOI=10.1073/pnas.242603899;
RA Strausberg R.L., Feingold E.A., Grouse L.H., Derge J.G., Klauener R.D., Collins B., Wagner L., Shenmen C.M., Schuler G.D., Altschul S.F., Zeeberg B., Buetow K.H., Schaefer C.F., Bhat N.K., Hopkins R.F., Jordan H., Moore T., Max S.I., Wang J., Heien F., Diatchenko L., Marusha K., Farmer A.A., Rubin G.M., Hong L., Stempelman M., Soares M.B., Bonaldo M.F., Casavant T.L., Scheetz T.E., Brownstein M.J., Udell T.B., Toshiyuki S., Carninci P., Muliyil S., Raha S.S., Loggellano N.A., Peters G.J., Abramson R.D., Mullaly S.J., Bosak S.A., McEwan P.J., McKernan K.J., Malek J.A., Gutarra P.H., Richards S., Worley K.C., Hale S., Garcia A.M., Gay L.V., Huiyk R.A., Vallaloun D.K., Muzny D.M., Sodergren E.J., Lu X., Gibbs R.A., Pahey J., Helton E., Kettman M., Madan A., Rodriguez S., Sanchez A., Whitting M., Madan A., Young A.C., Shevchenko Y., Bouffard G.G., Blakeley R.W., Touchman J.W., Green E.D., Dickson M.C., Rodriguez A.C., Grimwood J., Schmutz J., Myers R.M., Butcherfield Y.S.N., Krzywinski M.I., Skalska U., Smalme D.E., Schnerbach A., Schein J.B., Jones S.J.M., Marra M.A.;
RT "Generation and initial analysis of more than 15,000 full-length human and mouse cDNA sequences.";

Proc. Natl. Acad. Sci. U.S.A. 99:16899-16903(2002).
 CC -1- FUNCTION: Orphan receptor. Probably involved in the function of nociceptive neurons. May regulate nociceptor function and/or development, including the sensation or modulation of pain.
 CC Potentially activated by enkephalins (By similarity).
 CC -1- SUBCELLULAR LOCATION: Integral membrane protein.
 CC -1- TISSUE SPECIFICITY: Uniquely localized in a subset of small dorsal root and trigeminal sensory neurons.
 CC -1- SIMILARITY: Belongs to family 1 of G-protein coupled receptors. Mas subfamily.
 CC -----
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration between the Swiss Institute of Bioinformatics and the EMBL outstation at the European Bioinformatics Institute. There are no restrictions on its use by non-profit institutions as long as its content is in no way modified and this statement is not removed. Usage by and for commercial entities requires a license agreement (See <http://www.isb-sib.ch/announce/> or send an email to license@isb-sib.ch).
 CC -----
 DR EMBL: AY042215; AAK91806.1; -
 DR EMBL: AF474987; AAL86878.2; -
 DR EMBL: BC067292; AAH67292.1; -
 DR MIM: 607229; -
 DR InterPro: IPR000276; GPCR_Rhodopsn.
 DR Pfam: PF00001; 7tm_1; 1.
 DR PRINTS: PR00237; GPCR_RHODOPSIN.
 DR PROSITE: PS00237; G-PROTEIN_RECEP_F1_1; 1.
 DR PROSITE: PS50262; G-PROTEIN_RECEP_F1_2; 1.
 KW G-protein coupled receptor; Polymorphism; Transmembrane.
 FT DOMAIN 1 31 Extracellular (Potential).
 FT TRANSSEM 32 52 1 (Potential).
 FT TRANSSEM 53 60 Cytoplasmic (Potential).
 FT TRANSSEM 61 81 2 (Potential).
 FT TRANSSEM 82 96 Extracellular (Potential).
 FT TRANSSEM 97 117 3 (Potential).
 FT TRANSSEM 118 140 Cytoplasmic (Potential).
 FT TRANSSEM 141 161 4 (Potential).
 FT TRANSSEM 162 177 Extracellular (Potential).
 FT TRANSSEM 178 198 5 (Potential).
 FT TRANSSEM 199 213 Cytoplasmic (Potential).
 FT TRANSSEM 214 234 6 (Potential).
 FT TRANSSEM 235 254 Extracellular (Potential).
 FT TRANSSEM 255 275 7 (Potential).
 FT TRANSSEM 276 322 Cytoplasmic (Potential).
 FT VARIANT 169 169 D -> N (in dbSNP:4274188).
 FT VARIANT /FTID=VAR 019434.
 FT CONFLICT 3 3 S -> P (in Ref. 2).
 FT CONFLICT 82 82 C -> R (in Ref. 3).
 FT CONFLICT 307 307 W -> Q (in Ref. 3).
 FT CONFLICT 319 319 R -> K (in Ref. 2).
 KW SEQUENCE 322 AA; 36484 MW; 2531B8F0CB4EB74 CRC64;
 SQ
 Query Match 95.7%; Score 1688; DB 1; Length 322;
 Best Local Similarity 100.0%; Pred. No. 1.2e-115;
 Matches 322; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

DB 241 RHLDWKVLPCHVHLVSLFSLANSSANPFIYFVGSPFRONONKLVLPALQDTPB 300
 QY 316 VDEGGWLPQETLELSSRLBO 337
 DB 301 VDEGGWLPQETLELSSRLBO 322
 RESULT 2
 AAH67292
 ID AAH67292 PRELIMINARY; PRT; 322 AA.
 AC AAH67292;
 DT 25-MAR-2004 (TRENDEL. 27, Created)
 DT 25-MAR-2004 (TRENDEL. 27, Last sequence update)
 DT 25-MAR-2004 (TRENDEL. 27, Last annotation update)
 DE G protein-coupled receptor MRG3.
 GN MRG3.
 OS Homo sapiens (Human).
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 OC Mammalia; Eutheria; Primates; Catarrhini; Homiidae; Homo.
 OX NCBI_TaxID=9606;
 RX [1]
 RP SEQUENCE FROM N.A.
 RC TISSUE=Testis;
 RC MEDLINE=22388257; PubMed=12477932;
 RA Strausberg R.L., Feingold E.A., Grouse L.H., Derge J.G.,
 RA Klausner R.D., Collins F.S., Wagner C.M., Schuler G.D.,
 RA Altschul S.F., Zeeberg B., Buetow K.H., Schaefer C.F., Bhat N.K.,
 RA Hopkins R.F., Jordan H., Moore T., Max S.I., Wang J., Hsieh F.,
 RA Diatchenko L., Marusina K., Farmer A.A., Rubin G.M., Hong L.,
 RA Stapleton M., Soares M.B., Bonaldo M.F., Casavant T.L., Scheetz T.E.,
 RA Brownstein M.J., Uebli T.B., Toshyluk S., Carrinot P., Prange C.,
 RA Raha S.S., Loguella N.A., Peters G.J., Abramson R.D., Mullany S.J.,
 RA Bosak S.A., McKernan P.J., McKernan K.J., Malek J.A., Gunaratne P.H.,
 RA Richards S., Worley K.C., Hale S., Garcia A.M., Gay L.J., Hulyk S.W.,
 RA Villalón D.K., Mazny D.M., Sodergren E.J., Lu X., Gibbs R.A.,
 RA Fahney J., Helton E., Kettelman M., Madan A., Rodriguez Y., Bouffard G.,
 RA Whiting M., Madan A., Young A.C., Shevchenko Y., Bouffard G.,
 RA Blakesley R.W., Touchman J.W., Green E.D., Dickson M.C.,
 RA Rodriguez A.C., Grimwood J., Schmutz J., Myers R.M., Butcherfield Y.S.,
 RA Krzywicki M.I., Skalska U., Smalins D.E., Scherch A., Schein J.E.,
 RA Jones S.J., Marra M.A.;
 RT "Generation and initial analysis of more than 15,000 full-length human
 RT and mouse cDNA sequences.";
 RL Proc. Natl. Acad. Sci. U.S.A. 99:16899-16903(2002).
 RN [2]
 RP SEQUENCE FROM N.A.
 RC TISSUE=Testis;
 RC Strausberg R.;
 RL Submitted (MAR-2004) to the EMBL/GenBank/DBJ databases.
 DR EMBL: BC067292; AAH67292.1; -
 KW Receptor.
 SQ SEQUENCE 322 AA; 36479 MW; B91DC082B6D95DA8 CRC64;
 Query Match 94.3%; Score 1663; DB 2; Length 322;
 Best Local Similarity 99.4%; Pred. No. 7.9e-114;
 Matches 320; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

QY 256 RIHLDMKVLFCYHVLVIFLSALNSSANPIYFVGSFRORONRMLKVLQALADTPE 315
 DB 241 RIHLDMKVLFCYHVLVIFLSALNSSANPIYFVGSFRORONRMLKVLQALADTPE 300
 QY 316 VDEGGMLPQETLELSSRLQ 337
 DB 301 VDEGGMLPQETLELSSRLQ 322

RESULT 3
 SNS2_HUMAN STANDARD; PRT; 322 AA.
 AC OETD9;
 DT 01-OCT-2004 (Rel. 45, Created)
 DT 01-OCT-2004 (Rel. 45, Last sequence update)
 DT 01-OCT-2004 (Rel. 45, Last annotation update)
 DE Sensory neuron-specific G-protein coupled receptor 2.
 GN Name=SNRS2;
 OS Homo sapiens (human).
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 OC Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.
 OX NCBI_TaxID=9606;
 [1]
 SEQUENCE FROM N.A. AND TISSUE SPECIFICITY.
 RX MEDLINE=21853733; PubMed=11850634; DOI=10.1038/n815;
 RA Lembo P.M.C., Grazzini E., Groblewski T., O'Donnell D., Roy M.-O.,
 Zhang J., Hoffert C., Cao J., Schmidt R., Pelletier M., Labarre M.,
 Gosselin M., Fortin Y., Banville D., Shen S., Stroem P., Payza K.,
 Dray A., Walker P., Ahmad S.;
 RT "Proenkephalin A gene products activate a new family of sensory
 neuron-specific GPCRs.";
 RL Nat. Neurosci. 5:201-209(2002).
 CC -1- FUNCTION: Orphan receptor. Probably involved in the function of
 nociceptive neurons. May regulate nociceptor function and/or
 development, including the sensation or modulation of pain.
 CC Potentially activated by enkephalins (By similarity).
 CC -1- SUBCELLULAR LOCATION: Integral membrane protein.
 CC -1- TISSUE SPECIFICITY: Uniquely localized in a subset of small dorsal
 root and trigeminal sensory neurons.
 CC -1- SIMILARITY: Belongs to family 1 of G-protein coupled receptors.
 CC Mas subfamily.
 CC -----
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 between the Swiss Institute of Bioinformatics and the EMBL Outstation-
 CC the European Bioinformatics Institute. There are no restrictions on its
 CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement (See <http://www.ebi.ac.uk/announcements>
 CC or send an email to license@ebi.ac.uk).
 CC -----
 DR EMBL: AF749488; AAL68679.2; -
 DR InterPro: IPR000276; GPCR_Rhodopsin.
 DR Pfam: PFO0001; 7tm.1; 1.
 DR PRINTS: PR00237; GPCRHOOPS.
 DR PROSITE: PS00237; G_PROTEIN_RECEP_F1_1; 1.
 DR PROSITE: PS50262; G_PROTEIN_RECEP_F1_2; 1.
 KW G-protein coupled receptor; Transmembrane.
 FT DOMAIN 1 32
 FT TRANSMEM 33 53
 FT DOMAIN 54 60
 FT TRANSMEM 61 81
 FT DOMAIN 82 96
 FT TRANSMEM 97 117
 FT DOMAIN 118 140
 FT TRANSMEM 141 161
 FT DOMAIN 162 177
 FT TRANSMEM 178 198
 FT DOMAIN 199 213
 FT TRANSMEM 214 234
 FT DOMAIN 235 254
 FT TRANSMEM 255 275
 FT DOMAIN 276 322

FT CARBOHYD 89 89 N-linked (GlcNAc...) (Potential).
 SQ SEQUENCE - 322 AA; 36594 MW; D8C24308E3B4611B CRC64;
 Query Match 91.1%; Score 1606; DB 1; Length 322;
 Best Local Similarity 96.0%; Pred. No. 1.2e-109;
 Matches 309; Conservative 4; Mismatches 9; Indels 0; Gaps 0;

QY 16 MDSTIPVLTGELTPINGREETPCYKOTLSTFTGTCIVSLVALGNNAVLMLCGRMRNA 75
 DB 1 MDPTVPVLTGELTPINGREETPCYKOTLSTFTGTCIVSLVALGNNAVLMLCGRMRNA 60
 QY 76 VSIYIINLVADPLFLSGHICSPRLINIRHPSKILSPVMTPEPYIGLSAISTER 135
 DB 61 VSIYIINLVADPLFLSGHICSPRLINIRHPSKILSPVMTPEPYIGLSAISTER 120
 QY 136 CUSIIMPVYHCRPRYLSSVCMVLLMALSLSLSMWCDPLFGSADSVKETSPT 195
 DB 121 CUSIIMPVYHCRPRYLSSVCMVLLMALSLSLSMWCDPLFGSADSVKETSPT 180
 QY 196 IAWLVFVLCVVLGGSSVLVLRILGSRKMPVLTLYTILTVLVFLCGPFGIOMALFS 255
 DB 181 IAWLVFVLCVVLGGSSVLVLRILGSRKMPVLTLYTILTVLVFLCGPFGIOMALFS 240
 QY 256 RIHLDMKVLFCYHVLVIFLSALNSSANPIYFVGSFRORONRMLKVLQALADTPE 315
 DB 241 RIHLDMKVLFCYHVLVIFLSALNSSANPIYFVGSFRORONRMLKVLQALADTPE 300
 QY 316 VDEGGMLPQETLELSSRLQ 337
 DB 301 VDEGGMLPQETLELSSRLQ 322

RESULT 4
 SNS3_HUMAN STANDARD; PRT; 322 AA.
 AC OETD9;
 DT 01-OCT-2004 (Rel. 45, Created)
 DT 01-OCT-2004 (Rel. 45, Last sequence update)
 DT 01-OCT-2004 (Rel. 45, Last annotation update)
 DE Sensory neuron-specific G-protein coupled receptor 3.
 GN Name=SNRS3;
 OS Homo sapiens (human).
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 OC Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.
 OX NCBI_TaxID=9606;
 [1]
 SEQUENCE FROM N.A. AND TISSUE SPECIFICITY.
 RX MEDLINE=21853733; PubMed=11850634; DOI=10.1038/n815;
 RA Lembo P.M.C., Grazzini E., Groblewski T., O'Donnell D., Roy M.-O.,
 Zhang J., Hoffert C., Cao J., Schmidt R., Pelletier M., Labarre M.,
 Gosselin M., Fortin Y., Banville D., Shen S., Stroem P., Payza K.,
 Dray A., Walker P., Ahmad S.;
 RT "Proenkephalin A gene products activate a new family of sensory
 neuron-specific GPCRs.";
 RL Nat. Neurosci. 5:201-209(2002).
 CC -1- FUNCTION: Orphan receptor. Probably involved in the function of
 nociceptive neurons. May regulate nociceptor function and/or
 development, including the sensation or modulation of pain.
 CC Potentially activated by enkephalins including BAM2 (bovine adrenal
 medulla peptide 22) and BAM (8-22). BAM2 is the most potent
 compound and evoked a large and dose-dependent release of
 intracellular calcium in stably transfected cells. G(alpha)q
 proteins are involved in the calcium-signaling pathway.
 CC -1- SUBCELLULAR LOCATION: Integral membrane protein.
 CC -1- TISSUE SPECIFICITY: Uniquely localized in a subset of small dorsal
 root and trigeminal sensory neurons.
 CC -1- SIMILARITY: Belongs to family 1 of G-protein coupled receptors.
 CC Mas subfamily.
 CC -----
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 between the Swiss Institute of Bioinformatics and the EMBL Outstation-
 CC the European Bioinformatics Institute. There are no restrictions on its
 CC use by non-profit institutions as long as its content is in no way

CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
 CC or send an email to license@isb-sib.ch).

DR EMBL, AF474989; AAL86880.2; -
 DR InterPro; IPR000276; GPCR_Rhodopsn.
 DR Pfam; PF00001; 7tm_1; 1.
 DR PRINTS; PR00237; GPCR_RHODOPSIN.
 DR PROSITE; PS00237; G_PROTEIN_RECP_F1_1; 1.
 DR PROSITE; PS50262; G_PROTEIN_RECP_F1_2; 1.
 KM G-protein coupled receptor; Glycoprotein, Transmembrane.
 FT DOMAIN 1 31
 FT TRANSMEM 32 52
 FT DOMAIN 53 67
 FT TRANSMEM 68 88
 FT DOMAIN 89 96
 FT TRANSMEM 97 117
 FT DOMAIN 118 144
 FT TRANSMEM 145 165
 FT DOMAIN 166 177
 FT TRANSMEM 178 198
 FT DOMAIN 199 221
 FT TRANSMEM 222 242
 FT DOMAIN 243 254
 FT TRANSMEM 255 275
 FT DOMAIN 276 322
 FT CARBOHYD 16
 SQ SEQUENCE 322 AA; 36287 MW; 4C43B33E5DCEBFF5 CRC64;
 Query Match 80.0%; Score 1411; DB 1; Length 322;
 Best Local Similarity 84.5%; Pred. No. 2.2e-95;
 Matches 272; Conservative 16; Mismatches 34; Indels 0; Gaps 0;

QY 16 MDSTIPVLTGTELTPTNGREETPCVQKTLSPFGTCTIVSVLTGNAVYLMILGCMRRNA 75
 DB 1 MDPTVSTLDTLTLTNGTEETCTKQTLSTLVTCIVSVLTGNAVYLMILGCMRRNA 60
 QY 76 VSIYIILNVADFLFSGHIIICSPRLINIRHPISKILSPWTPPYFICGLSMISAISTER 135
 DB 61 FSIYIILNVADFLFSGHIIICSPRLINIRHPISKILSPWTPPYFICGLSMISAISTER 120
 QY 136 CLSLIMPIWHRCHPRRISSVWCYLMALSLRSLILEMFCDFLFGADSWCERSDPT 195
 DB 121 CLSLIMPIWHRCHPRRISSVWCYLMALSLRSLILEMFCDFLFGADSWCERSDPT 180
 QY 196 IAWLVPCVVLGSSSLVTLVRLICGRKMPRLVYTLITLVFLVLCGLPGIOMALFS 255
 DB 181 VAWLIFCVVLGSSSLVTLVRLICGRKMPRLVYTLITLVFLVLCGLPGIOMALFS 240
 QY 256 RIHLDMKVLFCVHLVSIPLSALNSSANPIIYFVGSPFRORONRLKVLQALQDTP 315
 DB 241 MIHDMKVLFCVHLVSIPLSALNSSANPIIYFVGSPFRORONRLKVLQALQDTP 300
 QY 316 VDEGGWLPQETLISGRLEQ 337
 DB 301 VDEGGWLPQETLISGRLEQ 322
 RESULT 5
 MREGI_HUMAN STANDARD; PRT; 322 AA.
 AC 096LB2; Q8TDB8; Rel. 45, Created)
 DT 01-OCT-2004 (Rel. 45, Last sequence update)
 DT 01-OCT-2004 (Rel. 45, Last annotation update)
 DE Mas-related G-protein coupled receptor member X1 (Sensory neuron-
 specific G-protein coupled receptor 4).
 GN Name=MREGX1; Synonyms=SNSR4;
 OS Homo sapiens (Human).
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 OC Mammalia; Eutheria; Primates; Catarrhini; Homidae; Homo.
 NC NCBL_TaxID=9606;
 RN [1]

RP SEQUENCE FROM N.A.
 RX MEDLINE=2435808; PubMed=11551509;
 RA Dong X., Han S.-K., Zylika M.J., Simon M.I., Anderson D.J.;
 RT "A diverse family of GPCRs expressed in specific subsets of
 RT nociceptive sensory neurons."
 RL Cell 106:619-632(2001).
 RN [2]
 RP SEQUENCE FROM N.A., AND TISSUE SPECIFICITY.
 RX MEDLINE=21853733; PubMed=11850634; DOI=10.1038/nrn15;
 RA Lembo P.M.C., Grazzini E., Groblewski T., O'Donnell D., Roy M.-O.,
 RA Zhang J., Hoffert C., Cao J., Schmidt R., Pelletier M., Labarre M.,
 RA Gosselin M., Fortin Y., Barville D., Shen S., Strom P., Payza K.,
 RA Dray A., Walker P., Ahmad S.;
 RL "Proenkephalin A gene products activate a new family of sensory
 RL neuron-specific GPCRs."
 RL Nat. Neurosci. 5:201-209(2002).
 RN [3]
 RP SEQUENCE FROM N.A.
 RX MEDLINE=22040266; PubMed=12044878;
 RA Takeda S., Kadowaki S., Haga T., Takeasu H., Mitaku S.;
 RL "Identification of G protein-coupled receptor genes from the human
 RL genome sequence."
 RL FEBS Lett. 520:97-101(2002).
 RN [4]
 RP SEQUENCE FROM N.A.
 RA Suwa M., Sato T., Okouchi I., Arita M., Futami K., Matsumoto S.,
 RA Tsumutsu S., Aburatani H., Asai K., Akiyama Y.;
 RL "Genome-wide discovery and analysis of human seven transmembrane helix
 RL receptor genes."
 RL Submitted (JUL-2001) to the EMBL/GenBank/DBJ databases.
 CC -1- FUNCTION: Orphan receptor. Probably involved in the function of
 CC nociceptive neurons. May regulate nociceptor function and/or
 CC development, including the sensation or modulation of pain.
 CC Potentially activated by enkephalins including BAM22 (bovine adrenal
 CC medulla peptide 22) and BAM (8-22). BAM22 is the most potent
 CC compound and evoked a large and dose-dependent release of
 CC intracellular calcium in stably transfected cells. G(alpha)q
 CC proteins are involved in the calcium-signaling pathway.
 CC -1- SUBCELLULAR LOCATION: Integral membrane protein.
 CC -1- TISSUE SPECIFICITY: Uniquely localized in a subset of small dorsal
 CC root and trigeminal sensory neurons.
 CC -1- SIMILARITY: Belongs to family 1 of G-protein coupled receptors.
 CC Mas subfamily.
 CC
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
 CC the European Bioinformatics Institute. There are no restrictions on its
 CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
 CC or send an email to license@isb-sib.ch).

DR EMBL, AY042213; AAK91804.1; -
 DR EMBL, AF474989; AAL86880.1; -
 DR EMBL, AB083628; BAB89341.1; -
 DR EMBL, AB065846; BAC06064.1; -
 DR MIM; 607227; -
 DR InterPro; IPR000276; GPCR_Rhodopsn.
 DR Pfam; PF00001; 7tm_1; 1.
 DR PRINTS; PR00237; GPCR_RHODOPSIN.
 DR PROSITE; PS00237; G_PROTEIN_RECP_F1_1; 1.
 DR PROSITE; PS50262; G_PROTEIN_RECP_F1_2; 1.
 KM G-protein coupled receptor; Glycoprotein, Polymorphism; Transmembrane.
 FT DOMAIN 1 31
 FT TRANSMEM 32 52
 FT DOMAIN 53 67
 FT TRANSMEM 68 88
 FT DOMAIN 89 96
 FT TRANSMEM 97 117
 FT DOMAIN 118 144
 FT TRANSMEM 145 165
 FT DOMAIN 166 177
 FT TRANSMEM 178 198

```

CC FT DOMAIN 199 221 Cytoplasmic (Potential).
CC FT TRANSMEM 222 242 6 (Potential).
CC FT DOMAIN 243 254 Extracellular (Potential).
CC FT TRANSMEM 255 275 7 (Potential).
CC FT DOMAIN 276 322 Cytoplasmic (Potential).
CC FT CARBOHYD 16 16 N-linked (GLCNAC...) (Potential).
CC FT VARIANT 36 36 I -> V (in dbSNP:11024885).
CC FT CONFLICT 5 5 /P1Id=VAR_019432.
CC FT SEQUENCE 322 AA; 36250 MW; C7F3A9F4418B8AD1 CRC64;
Query Match 78.4%; Score 1383; DB 1; Length 322;
Best Local Similarity 83.5%; Pred. No. 2,4e-93;
Matches 269; Conservative 16; Mismatches 37; Indels 0; Gaps 0;

QY 16 MDSTIPVLGTELPINGREESTPCYKOTLSFTGLTCTIVSLVATGNAVAVMLGCRMRNA 75
DB 1 MDPTISTLDLPINGREESTPCYKOTLSFTGLTCTIVSLVATGNAVAVMLGCRMRNA 60
QY 76 VSTIYIINLVADLPFLSGHICSPRLINIRHPSKILSVMPTEPYFGISMLSAISTER 135
DB 61 FSTIYIINLVADLPFLSGHICSPRLINIRHPSKILSVMPTEPYFGISMLSAISTER 120
QY 136 CSTIIMPWYHCRPRYLSVWCVLWALSLRSILEMFCDFLFGSADSVWCETSDPIT 195
DB 121 CSTIIMPWYHCRPRYLSVWCVLWALSLRSILEMFCDFLFGSADSVWCETSDPIT 180
QY 196 IANLVPLCVVLCSSSLVLRILCGSRKMPRLTYTILLTVLVFLCGLPFGIQMALFS 255
DB 181 VAMLVPLCVVLCSSSLVLRILCGSRKMPRLTYTILLTVLVFLCGLPFGIQMALFS 240
QY 256 RIHLDMKVLFCGHVLYSIFLSALNSSANPITYFVGSFRORORONKLVQLQRLQDTP 315
DB 241 WIHVDEEVLFCHVLYSIFLSALNSSANPITYFVGSFRORORONKLVQLQRLQDTP 300
QY 316 VDEGGGMLPOETLELSSGRLEQ 337
DB 301 VDEGGGMLPOETLELSSGRLEQ 322

RESULT 6
SNRS HUMAN
ID SNRS HUMAN STANDARD; PRT; 322 AA.
AC 08TDD7;
DT 01-OCT-2004 (Rel. 45, Created)
DT 01-OCT-2004 (Rel. 45, Last sequence update)
DE 01-OCT-2004 (Rel. 45, Last annotation update)
DE Sensory neuron-specific G-protein coupled receptor 5.
GN Name=SNRS;
OS Homo sapiens (Human).
OC Eukaryota; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.
OC NCBI_TaxId=9606;
RN [1]
RP SEQUENCE FROM N.A., AND TISSUE SPECIFICITY.
RX MEDLINE=21853733; PubMed=11850634; DOI=10.1038/m815;
RX Lembo P.M.C., Grazzini E., Groblewski T., O'Donnell D., Roy M.-O.,
RX Zhang J., Hoffert C., Cao J., Schmidt R., Pelletier M., Labarre M.,
RX Gosselin M., Fortin Y., Banville D., Shen S., Stroem P., Payza K.,
RX Dray A., Walker P., Ahmad S.;
RA "Proenkephalin A gene products activate a new family of sensory
RT neuron-specific GPCRs."
RL Nat. Neurosci. 5:201-209(2002).
RL -1- FUNCTION: Orphan receptor. Probably involved in the function of
CC nociceptive neurons. May regulate nociceptor function and/or
CC development, including the sensation or modulation of pain.
CC -1- SUBCELLULAR LOCATION: Integral membrane protein.
CC -1- TISSUE SPECIFICITY: Uniquely localized in a subset of small dorsal
CC root and trigeminal sensory neurons.
CC -1- SIMILARITY: Belongs to family 1 of G-protein coupled receptors.
CC Mas subfamily.

```

```

CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.ebi.ac.uk/announcements
CC or send an email to license@ebi.ac.uk).
CC -----
CC DR EMBL, AF474991; AAL86882.1; -
CC DR InterPro; IPR000276; GPCR_Rhodopsin.
CC DR Pfam; PF00001; 7tm_1; 1.
CC DR PRINTS; PR00237; GPCR_Rhodopsin.
CC DR PROSITE; PS00237; G_PROTEIN_RECP_F1_1; 1.
CC DR PROSITE; PS00262; G_PROTEIN_RECP_F2_1; 1.
CC KM G-protein coupled receptor; Transmembrane.
CC FT DOMAIN 1 31 Extracellular (Potential).
CC FT TRANSMEM 32 52 1 (Potential).
CC FT DOMAIN 53 60 Cytoplasmic (Potential).
CC FT TRANSMEM 61 81 2 (Potential).
CC FT TRANSMEM 82 96 Extracellular (Potential).
CC FT TRANSMEM 97 117 3 (Potential).
CC FT DOMAIN 118 137 Cytoplasmic (Potential).
CC FT TRANSMEM 138 158 4 (Potential).
CC FT DOMAIN 159 177 Extracellular (Potential).
CC FT TRANSMEM 178 198 5 (Potential).
CC FT DOMAIN 199 218 Cytoplasmic (Potential).
CC FT TRANSMEM 219 239 6 (Potential).
CC FT DOMAIN 240 254 Extracellular (Potential).
CC FT TRANSMEM 255 275 7 (Potential).
CC FT DOMAIN 276 322 Cytoplasmic (Potential).
CC FT CARBOHYD 89 89 N-linked (GLCNAC...) (Potential).
CC SQ SEQUENCE 322 AA; 36423 MW; 3D6FBA5DDF9D0 CRC64;
Query Match 78.4%; Score 1382; DB 1; Length 322;
Best Local Similarity 83.4%; Pred. No. 2,9e-93;
Matches 268; Conservative 20; Mismatches 32; Indels 0; Gaps 0;

QY 16 MDSTIPVLGTELPINGREESTPCYKOTLSFTGLTCTIVSLVATGNAVAVMLGCRMRNA 75
DB 1 MDPTISTLDLPINGREESTPCYKOTLSFTGLTCTIVSLVATGNAVAVMLGCRMRNA 60
QY 76 VSTIYIINLVADLPFLSGHICSPRLINIRHPSKILSVMPTEPYFGISMLSAISTER 135
DB 61 VSTIYIINLVADLPFLSGHICSPRLINIRHPSKILSVMPTEPYFGISMLSAISTER 120
QY 136 CSTIIMPWYHCRPRYLSVWCVLWALSLRSILEMFCDFLFGSADSVWCETSDPIT 195
DB 121 CSTIIMPWYHCRPRYLSVWCVLWALSLRSILEMFCDFLFGSADSVWCETSDPIT 180
QY 196 IANLVPLCVVLCSSSLVLRILCGSRKMPRLTYTILLTVLVFLCGLPFGIQMALFS 255
DB 181 VAMLVPLCVVLCSSSLVLRILCGSRKMPRLTYTILLTVLVFLCGLPFGIQMALFS 240
QY 256 RIHLDMKVLFCGHVLYSIFLSALNSSANPITYFVGSFRORORONKLVQLQRLQDTP 315
DB 241 RIHLDMKVLFCGHVLYSIFLSALNSSANPITYFVGSFRORORONKLVQLQRLQDTP 300
QY 316 VDEGGGMLPOETLELSSGRLEQ 335
DB 301 VDEGGGMLPOETLELSSGRLEQ 320

RESULT 7
MRG4 HUMAN
ID MRG4 HUMAN STANDARD; PRT; 322 AA.
AC 09LAD9; 08TDD6;
DT 01-OCT-2004 (Rel. 45, Created)
DT 01-OCT-2004 (Rel. 45, Last sequence update)
DT 01-OCT-2004 (Rel. 45, Last annotation update)
DE Mas-related G-protein coupled receptor member X4 (Sensory neuron-
DE specific G-protein coupled receptor 6).
GN Name=MRG4; Synonyms=SNRS6;
OS Homo sapiens (Human).

```

OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 OC Mammalia; Eutheria; Primates; Catarrhini; Homiidae; Homo.
 OC NCBI_Taxid=9606;
 RN [1]
 RP SEQUENCE FROM N.A.
 RX MEDLINE=21435808; PubMed=11551509;
 RA Dong X., Han S.-K., Zylka M.J., Simon M.I., Anderson D.J.;
 RT "A diverse family of GPCRs expressed in specific subsets of
 nociceptive sensory neurons."
 RL Cell 106:619-632(2001).
 RN [2]
 RP SEQUENCE FROM N.A., AND TISSUE SPECIFICITY.
 RX MEDLINE=21853733; PubMed=11850634; DOI=10.1038/nrn815;
 RA Lembo P.M.C., Grazzini E., Groblewski T., O'Donnell D., Roy M.-O.,
 RA Zhang J., Hoffert C., Cao J., Schmidt R., Pelletier M., Labaree M.,
 RA Gosselin M., Fortin Y., Banville D., Shen S., Strom P., Paya K.,
 RA Dray A., Walker P., Ahmad S.;
 RT "Proenkephalin A gene products activate a new family of sensory
 neuron-specific GPCRs."
 RL Nat. Neurosci. 5:201-209(2002).
 CC -1- FUNCTION: Orphan receptor. Probably involved in the function of
 nociceptive neurons. May regulate nociception function and/or
 development, including the sensation or modulation of pain.
 CC Potentially activated by enkephalins (By similarity).
 CC -1- SUBCELLULAR LOCATION: Integral membrane protein.
 CC -1- TISSUE SPECIFICITY: Uniquely localized in a subset of small dorsal
 root and trigeminal sensory neurons.
 CC -1- SIMILARITY: Belongs to family 1 of G-protein coupled receptors.
 CC Mas subfamily.
 CC -----
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 between the Swiss Institute of Bioinformatics and the EMBL outstation -
 the European Bioinformatics Institute. There are no restrictions on its
 use by non-profit institutions as long as its content is in no way
 modified and this statement is not removed. Usage by and for commercial
 entities requires a license agreement (see <http://www.isb-sib.ch/announce/>
 or send an email to license@isb-sib.ch).
 CC -----
 CC EMBL: AY042216; AAK91807.1; -
 CC EMBL: AF474992; AAL86883.1; -
 CC MIM: 607230; -
 DR InterPro: IPR00276; GPCR_Rhodopsn.
 DR Pfam: Pf00001; 7tm_1; 1.
 DR PRINTS: PR00237; GPCR_Rhodopsn.
 DR PROSITE: PS00237; G-PROTEIN RECEPTOR_F1_1; 1.
 DR PROSITE: PS50262; G-PROTEIN RECEPTOR_F1_2; 1.
 KM G-protein coupled receptor; Glycoprotein; Polymorphism; Transmembrane.
 FT DOMAIN 1 31
 FT TRANSMEM 32 52
 FT DOMAIN 53 60
 FT TRANSMEM 61 81
 FT DOMAIN 82 96
 FT TRANSMEM 97 117
 FT DOMAIN 118 137
 FT TRANSMEM 138 158
 FT DOMAIN 159 177
 FT TRANSMEM 178 198
 FT DOMAIN 199 218
 FT TRANSMEM 219 239
 FT DOMAIN 240 254
 FT TRANSMEM 255 275
 FT DOMAIN 276 322
 FT CARBOHYD 25 25
 FT CARBOHYD 89 89
 FT VARIANT 8 8
 FT VARIANT 25 25
 FT VARIANT 54 54
 FT VARIANT 83 83
 FT CONFLICT 182 182

FT CONFLICT 319 319 R -> K (in Ref. 2).
 SQ SEQUENCE 322 AA; 36434 MW; 7CA676F8BD390A31 CRC64;
 Query Match 77.5%; Score 1367; DB 1; Length 322;
 Best Local Similarity 83.1%; Pred. No. 3.6e-92;
 Matches 266; Conservative 20; Mismatches 34; Indels 0; Gaps 0;
 QY 16 MDSTIPVAGTEPLTPINGREERTPCYKQTLSPFTGLCTIVSLVATGNAAVLLCGRRNRNA 75
 DB 1 MDPTVPVGTGLTPINGREERTPCYNQTLSPFTGLCTIVSLVATGNAAVLLCGRRNRNA 60
 QY 76 VSIYILNVAADFLFLSHIICSPRLNINHPISKILSPMTPTPTGISMISAISTER 135
 DB 61 VSIYILNVAADFLFLSPQIIRSPRLINSHLRKILVSWPTPTPTGISMISAISTER 120
 QY 136 CISTLPIWYHCRPRRYSSVYVCLVNLSTLRISILEMPCDFPFGSADSWCETSPFIT 195
 DB 121 CISTLPIWYHCRPRRYSSVYVCLVNLSTLRISILEMPCDFPFGSADSWCETSPFIT 180
 QY 196 IAMLVFLCVLCSGLVLLVRLICGSRMPRLTYVITLTVLPFLCGLPFGIQLWLF 255
 DB 181 VAMLIFFLCVVLCSGLVLLVRLICGSRMPRLTYVITLTVLPFLCGLPFGIQLWLF 240
 QY 256 RIHLDWKVLPFGVHLVSIFFLSALNSANPITFFVGSFRRQRONQKLVLRALOTPE 315
 DB 241 RHHNLLEVLVYCHVYLVCMSSLSNSANPITFFVGSFRRQRONQKLVLRALOTPE 300
 QY 316 VDEGGGWLPORTLELSSGRL 335
 DB 301 VDKGGQLPRESLSSGRL 320
 RESULT 8
 MSG2_HUMAN
 ID MSG2_HUMAN STANDARD; PRT; 330 AA.
 AC Q96LB1;
 DT 01-OCT-2004 (Rel. 45, Created)
 DT 01-OCT-2004 (Rel. 45, Last sequence update)
 DT 01-OCT-2004 (Rel. 45, Last annotation update)
 DE Mas-related G-protein coupled receptor member X2.
 GN Name=MRGX2;
 OS Homo sapiens (Human).
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 OC Mammalia; Eutheria; Primates; Catarrhini; Homiidae; Homo.
 OC NCBI_Taxid=9606;
 RN [1]
 RP SEQUENCE FROM N.A.
 RX MEDLINE=21435808; PubMed=11551509;
 RA Dong X., Han S.-K., Zylka M.J., Simon M.I., Anderson D.J.;
 RT "A diverse family of GPCRs expressed in specific subsets of
 nociceptive sensory neurons."
 RL Cell 106:619-632(2001).
 RN [2]
 RP SEQUENCE FROM N.A.
 RX MEDLINE=22040266; PubMed=12044878;
 RA Takeda S., Kadowaki S., Haga T., Takeasu H., Mitaku S.;
 RT "Identification of G protein-coupled receptor genes from the human
 genome sequence."
 RL FEBS Lett. 520:97-101(2002).
 RN [3]
 RP SEQUENCE FROM N.A.
 RA Suwa M., Sato T., Okouchi I., Arita M., Futami K., Matsumoto S.,
 RA Tautsuni S., Aburatani H., Asai K., Akiyama Y.;
 RT "Genome-wide discovery and analysis of human seven transmembrane helix
 receptor genes."
 RL Submitted (JUL-2001) to the EMBL/GenBank/DBJ databases.
 RN [4]
 RP SEQUENCE FROM N.A.
 RC TISSUE=Brain;
 RX MEDLINE=22388257; PubMed=12477932; DOI=10.1073/pnas.242603899;
 RA Straube R.L., Feingold E.A., Grouse L.H., Derge J.G.,
 RA Klausner R.D., Collins F.S., Wagner L., Sherman C.M., Schuler G.D.,
 RA Altschul S.F., Zeeberg B., Buetow K.H., Schaefer C.F., Bhat N.K.,

RA Hopkins R.F., Jordan H., Moore T., Max S.I., Wang J., Heifetz F.,
RA Diatchenko L., Marusina K., Farmer A.A., Rubin G.M., Hong L.,
RA Stapleton M., Soares M.B., Bonaldo M.F., Casavant T.L., Schetz T.E.,
RA Brownstein M.J., Usdin T.B., Toshiyuki S., Carninci P., Prange C.,
RA Raha S.S., Loquellano N.A., Peters G.J., Abramson R.D., Mullaly S.J.,
RA Bosak S.A., McEwan P.J., McKernan K.J., Malek J.A., Gunaratne P.H.,
RA Richards S., Worley K.C., Hale S., Garcia A.M., Gay L.J., Hulyk S.W.,
RA Villalon D.K., Murry D.M., Sodergren E.J., Lu X., Gibbs R.A.,
RA Fahy J., Helton E., Kettelman M., Madan A., Rodriguez S., Sanchez A.,
RA Whiting M., Madan A., Young A.C., Shevchenko Y., Bouffard G.G.,
RA Blakeley R.W., Touchman J.W., Green E.D., Dickson M.C.,
RA Rodriguez A.C., Grimwood J., Schmutz J., Myers R.M.,
RA Butterfield Y.S.N., Krzywinski M.I., Skalek U., Smalhus D.E.,
RA Scherch A., Schein J.E., Jones S.J.M., Marra M.A.,
RT "Generation and initial analysis of more than 15,000 full-length human
RT and mouse cDNA sequences.";
RL Proc. Natl. Acad. Sci. U.S.A. 99:16899-16903(2002).
RN (5)
RP TISSUE SPECIFICITY, AND POSSIBLE FUNCTION.
RX PubMed=12915402; DOI=10.1074/jbc.M302456200;
RA Robas N., Mead E., Fiddock M.;
RT "MrgX2 is a high potency corticistatin receptor expressed in dorsal root
RT ganglion.";
CC CC
CC J. Biol. Chem. 278:44400-44404(2003).
CC -1- FUNCTION: Orphan receptor. Probably involved in the function of
CC nociceptive neurons. May regulate nociceptive function and/or
CC development, including the sensation or modulation of pain.
CC Corticistatin-14 seems to be a high potency ligand at this receptor.
CC Corticistatin has several biological functions including roles in
CC sleep regulation locomotor activity, and cortical function. In
CC receptor-expressing cells, corticistatin-stimulated increases in
CC intracellular Ca(2+) but had no effect on basal or forskolin-
CC stimulated cAMP levels, suggesting that this receptor is G(q)-
CC coupled.
CC -1- SUBCELLULAR LOCATION: Integral membrane protein.
CC -1- TISSUE SPECIFICITY: Has a limited expression profile, both
CC peripheral and within the central nervous system, with highest
CC levels in dorsal root ganglion.
CC -1- SIMILARITY: Belongs to family 1 of G-protein coupled receptors.
CC Was subfamily.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
CC or send an email to license@isb-sib.ch).
CC -----
CC EMBL, AY042214; AAK91805.1; -;
DR EMBL, AB083626; BAB89339.1; -;
DR EMBL, AB065811; BAC06030.1; -;
DR EMBL, BC063450; AAH63450.1; -;
DR MIM: 607228; -;
DR InterPro: IPR000276; GPCR_Rhodopsin.
DR Pfam: PF00001; 7tm_1, 1.
DR PRINTS: PR00237; GPCRHHODOPSIN.
DR PROSITE: PS00237; G_PROTEIN_RECP_F1_1; 1.
DR PROSITE: PS00262; G_PROTEIN_RECP_F1_2; 1.
KW G-protein coupled receptor; Polymorphism; Transmembrane.
FT DOMAIN 1 33
FT TRANSMEM 34 54
FT TRANSMEM 55 63
FT TRANSMEM 64 84
FT TRANSMEM 85 96
FT TRANSMEM 97 117
FT TRANSMEM 118 144
FT TRANSMEM 145 165
FT TRANSMEM 166 184
FT TRANSMEM 185 205
FT TRANSMEM 206 228
FT TRANSMEM 229 249
FT DOMAIN 250 264
FT Extracellular (Potential).
FT Extracellular (Potential).

FT TRANSMEM 265 285 7 (Potential).
FT DOMAIN 286 330 Cytoplasmic (Potential).
FT VARIANT 62 62 N->S (in dbSNP:10833049).
FT SEQUENCE 330 AA; 37099 MW; 0B328FD78B1DFE8E CRC64;
SQ
Query Match 56.1%; Score 989.5; DB 1; Length 330;
Best Local Similarity 62.1%; Pred. No. 1.5e-64;
Matches 208; Conservative 27; Mismatches 79; Indels 21; Gaps 5;
QY 16 MDSTIPVLTETLPINGREET---PCYKOTLSPTGLTIVSVALTGNVAVLMLCCRR 72
DB 1 MDPTTAMGTSTVTYVNGNDQALLLCKGKTLIPVFILITVALVGLVGNFVLLCFRRK 60
QY 73 RNNAVSYIINLVAAADLPFLSHII-----CSPLINIRHISKLSPVMTFPYF 122
DB 61 RNAFSVYVSLGADLPFLCFQINCLVYLSNFFCS---ISTNPF--SPFTVMCAVL 114
QY 123 IGLSMSTAISTRCLSLIPIWYHCRPRPYLSVWCVLMLATSLRSILEMPCDFLPSG 182
DB 115 AGLSMLSTWSTERCLSVLPWYRCRPRHLSAVCVLMLATSLRSILEGKCGFLFD 174
QY 183 ADSWCESTDPTITAMLVFLCVLCCSSLVLLVRLICGSRKMPRLVYTLITLVFL 242
DB 175 GDSGWCQTFDPTITAMLVFLFVWLCCSSLALVRLICGSRGPLEFRLVYTLITLVFL 234
QY 243 CGIPFGIQLALPSRIHLDMKVLFCFVHLVSLPLSANSANPIYFNGSFRQRORON- 301
DB 235 CGIPFGIQLFLIMWKDSDVFCFHPVSVVLSINSANPIYFVGSFRQWRLQOP 294
QY 302 -LKLVLVRLAOPDPEYDECGWLPQETLRLSGRL 335
DB 295 LKLVLVRLAOPDPEYDECGWLPQETLRLSGRL 329
RESULT 9
AAH63450 PRELIMINARY; PRT; 330 AA.
AC AAH63450; (TEMBLrel. 27, Created)
DT 02-MAR-2004 (TEMBLrel. 27, last sequence update)
DT 02-MAR-2004 (TEMBLrel. 27, last annotation update)
DE G protein-coupled receptor MKG2.
GN MKG2.
OS Homo sapiens (Human).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Primates; Catarrhini; Homiidae; Homo.
OX NCBI_TaxID=9606;
OX [1]
RP SEQUENCE FROM N.A.
RC TISSUE=Brain;
RX MEDLINE=22388257; PubMed=12477932;
RA Strausberg R.L., Feingold E.A., Grouse L.H., Derge J.G.,
RA Klausner R.D., Collins F.S., Wagner L., Shenmen C.M., Schuler G.D.,
RA Altschul S.F., Zeeberg B., Buettow K.H., Schaefer C.F., Bhat N.K.,
RA Hopkins R.F., Jordan H., Moore T., Max S.I., Wang J., Heifetz F.,
RA Diatchenko L., Marusina K., Farmer A.A., Rubin G.M., Hong L.,
RA Stapleton M., Soares M.B., Bonaldo M.F., Casavant T.L., Schetz T.E.,
RA Brownstein M.J., Usdin T.B., Toshiyuki S., Carninci P., Prange C.,
RA Raha S.S., Loquellano N.A., Peters G.J., Abramson R.D., Mullaly S.J.,
RA Bosak S.A., McEwan P.J., McKernan K.J., Malek J.A., Gunaratne P.H.,
RA Richards S., Worley K.C., Hale S., Garcia A.M., Gay L.J., Hulyk S.W.,
RA Villalon D.K., Murry D.M., Sodergren E.J., Lu X., Gibbs R.A.,
RA Fahy J., Helton E., Kettelman M., Madan A., Rodriguez S., Sanchez A.,
RA Whiting M., Madan A., Young A.C., Shevchenko Y., Bouffard G.G.,
RA Blakeley R.W., Touchman J.W., Green E.D., Dickson M.C.,
RA Rodriguez A.C., Grimwood J., Schmutz J., Myers R.M., Butterfield Y.S.,
RA Krzywinski M.I., Skalek U., Smalhus D.E., Scherch A., Schein J.E.,
RA Jones S.J., Marra M.A.,
RT "Generation and initial analysis of more than 15,000 full-length human
RT and mouse cDNA sequences.";
RL Proc. Natl. Acad. Sci. U.S.A. 99:16899-16903(2002).
RN [2]

RP SEQUENCE FROM N.A.
 RC TISSUE=Brain;
 RA Strausberg R.;
 RL Submitted (DEC-2003) to the EMBL/Genbank/DBJ databases.
 DR EMBL: BC63450; AAB63450.1; -
 KW Receptor.
 SQ SEQUENCE 330 AA; 37099 MW; 0B328FD78B1D66B CRC64;

Query Match 56.1%; Score 989.5; DB 2; Length 330;
 Best Local Similarity 62.1%; Pred. No. 1.5e-64;
 Matches 208; Conservative 27; Mismatches 79; Indels 21; Gaps 5;

QY 16 MDSTIPVLGTELPINGREET--PCYKQTLSTGLTCIVSLVALTGNAVVLMLGCR 72
 DB 1 MDPTTPAMGTESTVINGNDQALLLCGETLIPVFLIFLALVGLVNGFVLMILGFRMR 60
 QY 73 RNVSITILMLVADFLPLSGHIT-----CSPLRLINIRPIKSLSPWTFPYF 122
 DB 61 RNAPSVYVLISLAGDFLFCQIINCIVLYLSNPFCS---ISINPP-SFTVTWCAYL 114
 QY 123 IGLSMLSAISTERCLSTLPIWYHCRPRYLSVSWCVLLMALSLRSILEMFCDFLPSG 182
 DB 115 AGLSMLSTVSTERCLSTLPIWYHCRPRYLSVSWCVLLMALSLRSILEMFCDFLPSD 174
 QY 183 ADSVWCETSDPTITAMLVFLCGSSLVLLVRLICGSRKMPRLRYTTLITLVLPFL 242
 DB 175 GDSGMCCTPPTITAMLVFLFMVLCGSSLVLLVRLICGSRGLPLRLYTLITLVLPFL 234
 QY 243 CGLPFGIQLMFLSRILHDMKLVFCHVHVSIFLSALNSANPIITYFVGSFRQRORNLQ 301
 DB 235 CGLPFGIQLMFLSRILHDMKLVFCHVHVSIFLSALNSANPIITYFVGSFRQRORNLQ 294
 QY 302 -LKLVLQALADPTPEVDEGGGMLPOETTELSGSR 335
 DB 295 ILKLVLQALADPTPEVDEGGGMLPOETTELSGSR 329

RESULT 10
 Q7TN42 PRELIMINARY; PRT: 323 AA.
 ID Q7TN42
 AC Q7TN42; (TrEMBLrel. 25, Created)
 DT 01-OCT-2003 (TrEMBLrel. 25, Last sequence update)
 DT 01-OCT-2004 (TrEMBLrel. 26, Last annotation update)
 DE MRG G protein-coupled receptor.
 GN Name=Mrgc;
 OS Rattus norvegicus (Rat).
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Rattus.
 OX NCBI_TaxID=10116;
 RN [1]
 RP SEQUENCE FROM N.A.
 RC STRAIN=Sprague-Dawley;
 RX MEDLINE=22810130; PubMed=12909716;
 RA Zylka M.J., Dong X., Southwell A.L., Anderson D.J.;
 RT "Atypical expansion in mice of the sensory neuron-specific Mrg G
 RT protein-coupled receptor family."
 RL Proc. Natl. Acad. Sci. U.S.A. 100:10043-10048(2003).
 DR EMBL: AF518245; AA008317.1;
 DR GO: GO:0016021; C: integral to membrane, IEA.
 DR GO: GO:0004872; F: receptor activity, IEA.
 DR GO: GO:000184; F: rhodopsin-like receptor activity, IEA.
 DR GO: GO:0007186; P: G-protein coupled receptor protein signaling, IEA.
 DR InterPro: IPR000276; GPCR_Rhodpsn.
 DR Pfam: PF00001; 7tm.1.1.
 DR PROSITE: PS00237; G_PROTEIN_RECP_F1_1; UNKNOWN_1.
 DR PROSITE: PS50262; G_PROTEIN_RECP_F1_2; 1.
 KW Receptor.
 SQ SEQUENCE 323 AA; 37136 MW; 50639PDA42E8052 CRC64;

Query Match 46.7%; Score 823.5; DB 2; Length 323;
 Best Local Similarity 52.8%; Pred. No. 2.1e-52;
 Matches 172; Conservative 51; Mismatches 92; Indels 11; Gaps 6;

QY 16 MDSTIPVLGTELPINGREETPCYKQTLSTGLTCIVSLVALTGNAVVLMLGCRKRNA 75
 DB 1 MDPTTPAMGTESTVINGNDQALLLCGETLIPVFLIFLALVGLVNGFVLMILGFRMR 59
 QY 76 VSITILMLVADFLPLSGHITCSPLRLINT----RHPISK-IISPWTFPYFGLSMLSA 130
 DB 60 ISVYVLNLSLDSFPLCFIDSLMRIMNPFYIAHKLKRIKLGNAVPIFYISGLSLSA 119
 QY 131 ISTERCLSTLPIWYHCRPRYLSVSWCVLLMALSLRSILEMFCDFLPSGADVWCET 190
 DB 120 ISTERCLSTLPIWYHCRPRYLSVSWCVLLMALSLRSILEMFCDFLPSGADVWCET 178
 QY 191 SDPTITAMLVFLCGSSLVLLVRLICGSRKMPRLRYTTLITLVLPFLCGLPFGIQ 250
 DB 179 VDFITAVLFLFLMFLCGSSLVLLVRLICGSRKMPRLRYTTLITLVLPFLCGLPGLY 238
 QY 251 WAL--FSRILHDMKLVFCHVHVSIFLSALNSANPIITYFVGSFRQRORNLQKLVLR 308
 DB 239 LFLYVFGIHLHYR--FCHIVQTVLLSCVSSANPIITYLVGSFRQRORNLQKLVLR 296
 QY 309 ALDPTPEVDEGGGMLPOETTELSGSR 334
 DB 297 ALDPTPEVDEGGGMLPOETTELSGSR 322

RESULT 11
 SNS1_RAT STANDARD; PRT: 323 AA.
 ID SNS1_RAT
 AC OBRAG1.
 DT 01-OCT-2004 (Rel. 45, Created)
 DT 01-OCT-2004 (Rel. 45, Last sequence update)
 DT 01-OCT-2004 (Rel. 45, Last annotation update)
 DE Sensory neuron-specific G-protein coupled receptor 1.
 GN Name=Snrt1; Synonyms=Snrt;
 OS Rattus norvegicus (Rat).
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Rattus.
 OX NCBI_TaxID=10116;
 RN [1]
 RP SEQUENCE FROM N.A., AND TISSUE SPECIFICITY.
 RC STRAIN=Sprague-Dawley;
 RX MEDLINE=21853733; PubMed=11850634; DOI=10.1038/nr815;
 RA Lembo P.M.C., Grazzini E., Groblewski T., O'Donnell D., Roy M.-O.,
 RA Zhang J., Hoffert C., Cao J., Schmidt R., Pelletier M., Labarre M.,
 RA Gosselin M., Fortin Y., Banville D., Shen S., Stroem P., Payza K.,
 RA Dray A., Walker P., Ahmad S.;
 RT "Proenkephalin A gene products activate a new family of sensory
 RT neuron-specific GPCRs."
 RL Nat. Neurosci. 5:201-209(2002).
 CC -!- FUNCTION: Orphan receptor. Probably involved in the function of
 CC nociceptive neurons. May regulate nociceptor function and/or
 CC development, including the sensation or modulation of pain.
 CC Potentially activated by enkephalins (By similarity).
 CC -!- TISSUE SPECIFICITY: Integral membrane protein.
 CC -!- SUBCELLULAR LOCATION: Uniquely localized in a subset of small dorsal
 CC root and trigeminal sensory neurons. Associated preferentially
 CC with IB4 class of small-diameter somatosensory afferents (also
 CC known as nociceptors).
 CC -!- SIMILARITY: Belongs to family 1 of G-protein coupled receptors.
 CC Mas subfamily.
 CC -----
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
 CC the European Bioinformatics Institute. There are no restrictions on its
 CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement (See <http://www.ebi.ac.uk/announcements/>
 CC or send an email to license@ebi-sib.ch).
 CC -----
 DR EMBL: AF474986; AAB6877.2; ALT_INIT.
 DR RGD: 632284; Snrt1.
 DR InterPro: IPR000276; GPCR_Rhodpsn.

[illegible]

Proc. Natl. Acad. Sci. U.S.A. 99:14740-14745(2002).

[2]

SEQUENCE FROM N.A.

RC STRAIN=C57BL/6;

RA Han S.-K., Dong X., Hwang J.-I., Zylka M.J., Anderson D.J.,

Simon M.I.;

RL Submitted (SEP-2002) to the EMBL/GenBank/DBJ databases.

DR EMBL: AY152435; C: integral to membrane; IEA.

DR GO: GO:0016021; C: integral to membrane; IEA.

DR GO: GO:0004872; F: receptor activity; IEA.

DR GO: GO:0001584; F: rhodopsin-like receptor activity; IEA.

DR GO: GO:0007186; P: G-protein coupled receptor protein signalin. . .; IEA.

DR InterPro: IPR000276; GPCR_Rhodopsn.

DR Pfam: PF00001; 7tm.1; 1.

DR PRINTS: PR00237; GPCR_RHODOPSN.

DR PROSITE: PS00237; G_PROTEIN_RECEP_F1_1; UNKNOWN_1.

DR PROSITE: PS00262; G_PROTEIN_RECEP_F1_2; 1.

KW Receptor.

SO SEQUENCE 322 AA; 36809 MW; 1E03CDD54AF8BDC CRC64;

Query Match 45.2%; Score 797; DB 2; Length 322;

Best Local Similarity 53.2%; Pred. No. 1.8e-50;

Matches 174; Conservative 43; Mismatches 96; Indels 14; Gaps 8;

16 MDSITPVIGLETITINGREERPCVK-QTLSPTGTLCTIVSYVALTGNVAVYMLGCRMRN 74

1 MDPITSHDVESTYPLNTGHPNCPIITLSL-LVLTITVLGLAGNTIVYMLGFRMRK 58

75 AVSIYIINLVADPLFLSGHICSPRLINL---RHPISK-IISPVTPEPYIGLSMLS 129

59 AISTYIINLVADPLFLCCHHIDBLNIIIDPYGYARKLSGLIDGNAAIIPYISGLSILS 118

130 AISTERCISITMPYIWHCRPRYVSSVYCVLWALSLRSTLEMFCDPLFGADSWCE 189

119 AISTERCICVLPMPYWHCHPRNMASATICALIWLSTFLMGILDW-FSGFLGETHHNLM-K 176

190 TSDITITAMLVFLCVLGGSSIVLVYIILGSRKMPITRLVYITLLTVLVFLLGLPFI 249

177 NVDTITITAPLITFLPMLSSGLALLRLICGPRKPSRLIYVITVYVYLLICGLPGL 236

250 QWAL-FSRHLDKVLFCVHVLVSIPLSALNSANPIYFVYVSGFRQQRNMLKVLQ 307

237 YLFLYIFGVHLHP--FCHLYQTAVLVSCNNSANPIYFLVSGFRQHRHSLKVLK 294

308 PALQDTPVEVDEGGMLPQETLELSGR 334

295 RALQDTPBEDEVYTDSHLKTETLSER 321

RESULT 13

MGA_RAT

ID MGA_RAT STANDARD; PRT; 304 AA.

AC 07TN49;

DT 01-OCT-2004 (Rel. 45, Created)

DT 01-OCT-2004 (Rel. 45, Last sequence update)

DT 01-OCT-2004 (Rel. 45, Last annotation update)

DE Maa-related G-protein coupled receptor member A.

GN Name=Mrgpra; Synonyms=Mrga;

OS Rattus norvegicus (Rat).

OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;

OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Rattus.

NCBI_TaxID=10116;

[1]

SEQUENCE FROM N.A.

RC STRAIN=Sprague-Dawley;

RX MBLIN=22810130; PubMed=12909716; DOI=10.1073/pnas.1732949100;

RA Zylka M.J., Dong X., Southwell A.L., Anderson D.J.;

RT "Atypical expansion in mice of the sensory neuron-specific Mrg G

protein-coupled receptor family."

RT Proc. Natl. Acad. Sci. U.S.A. 100:10043-10048(2003).

CC -1- FUNCTION: Orphan receptor. May regulate nociception function and/or

development, including the sensation or modulation of pain.

CC -1- SUBCELLULAR LOCATION: Integral membrane protein.

CC -1- SIMILARITY: Belongs to family 1 of G-protein coupled receptors.
 CC Mas subfamily.

CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
 CC the European Bioinformatics Institute. There are no restrictions on its
 CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
 CC or send an email to license@isb-sib.ch).

DR EMBL: AF518238; AA008310.1; --
 DR RGD; 738050; Mrgpra.
 DR InterPro; IPR000276; GPCR_Rhodopsn.
 DR Pfam; PF00001; 7tm_1; 1.
 DR PROSITE; PS00237; G_PROTEIN_RECP_F1_1; 1.
 DR PROSITE; PS50262; G_PROTEIN_RECP_F1_2; 1.
 KM G-protein coupled receptor; Transmembrane.
 FT DOMAIN 1 17 Extracellular (Potential).
 FT TRANSMEM 18 38 1 (Potential).
 FT DOMAIN 39 46 Cytoplasmic (Potential).
 FT TRANSMEM 47 67 2 (Potential).
 FT DOMAIN 68 80 Extracellular (Potential).
 FT TRANSMEM 81 101 3 (Potential).
 FT DOMAIN 102 132 Cytoplasmic (Potential).
 FT TRANSMEM 133 153 4 (Potential).
 FT DOMAIN 154 167 Extracellular (Potential).
 FT TRANSMEM 168 188 5 (Potential).
 FT DOMAIN 189 211 Cytoplasmic (Potential).
 FT TRANSMEM 212 232 6 (Potential).
 FT DOMAIN 233 244 Extracellular (Potential).
 FT TRANSMEM 245 265 7 (Potential).
 FT DOMAIN 266 304 Cytoplasmic (Potential).
 SQ SEQUENCE 304 AA; 34334 MW; 4A820A4E0192E886 CRC64;

Query Match 44.4%; Score 782; DB 1; Length 304;
 Best Local Similarity 53.3%; Pred. No. 2,1e-49;
 Matches 176; Conservative 34; Mismatches 84; Indels 36; Gaps 8;

QY 16 MDSTIPVLGTELTPINGREPTCYKQTLSTFTGLCTIVS-LVALTGNAAVLTMLGCRMRN 74
 DB 1 MDKTIPT-----GSFNGRTLLP-----NLIIIGLVGLTGNAAVFWLLGFRLLARN 45
 QY 75 AVSIYIINLVAADEFLFSGHIIICSPRLNIRHP---ISKILSPWTPPYFGLSMLSAI 131
 DB 46 AFSVYIINLVAADEFLFCHIIIDSTLLKRSYNIIFLPCEFTVMWVPYIAGLSMLSAI 105
 QY 132 STERCLSILMPDIWYHCRPRYLLSSVMCVLWALSLRSILSEMFCDLFSGADSVW---- 187
 DB 106 STERCLSVVCPIWYRCRPRKHTSTVMCSAIVLSLICILNRYFCGL---DTYRKDN 161
 QY 188 -CETSDFTITAMLVFLCVLTCGSSLVLLVRLICGRKMPRLTYVITLTVLVFLCGLP 246
 DB 162 RCLASNFPTAACLIFFLVVLCISLALIVRSFCAGGRKRLRYATIMLVVFLCGLP 221
 QY 247 FGIQWALFSRIHLDKVLFCVHVLVSIFLSALNSANDPIYFVGSFRORONRQNLKVL 306
 DB 222 FGIHWFLLIWKIDYGRKAYGLYLAALVLTAVNSCANPIYFVGSFR-HQKQTLKVL 280
 QY 307 QRALQDTPETAE-----NTVEMSSSKVE 336
 DB 281 QRALQDTPETAE-----NTVEMSSSKVE 303
 RESULT 14
 QY 091YB7 PRELIMINARY; PRT; 331 AA.
 AC 091YB7;
 DT 01-DEC-2001 (Tremblrel. 19; Created)
 DT 01-DEC-2001 (Tremblrel. 26; Last sequence update)
 DT 01-MAR-2004 (Tremblrel. 26; Last annotation update)
 DE G-protein coupled receptor.
 GN Name=rc 56.1.3;

OS Rattus norvegicus (Rat).
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 CC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Rattus.
 OX NCBI_TaxID=10116;

RP SEQUENCE FROM N.A.
 RX MEDLINE=22080228; PubMed=12084918;
 RA Bender B., Bust A., Jurzak M., Langlois X., Baggerman G.,
 RA Verasaelle P., Ercken M., Guo H.Q., Wintmolders C.,
 RA Van den Wyngaert I., Van Oers I., Schoofs L., Luyten W.,
 RT "Characterization of an orphan G protein-coupled receptor localized in
 RT the dorsal root ganglia reveals adenine as a signalling molecule";
 RL Proc. Natl. Acad. Sci. U.S.A. 99:8573-8578(2002).
 DR EMBL; AJ311952; CAC84592.1; --
 DR GO; GO:0016021; C:Integral to membrane; IEA.
 DR GO; GO:0004872; F:receptor activity; IEA.
 DR GO; GO:0001584; F:rhodopsin-like receptor activity; IEA.
 DR GO; GO:0007186; P:G-protein coupled receptor protein signaln. . .; IEA.
 DR InterPro; IPR00276; GPCR_Rhodopsn.
 DR Pfam; PF00001; 7tm_1; 1.
 DR PRINTS; PR00237; GPCR_RHODOPSIN.
 DR PROSITE; PS00237; G_PROTEIN_RECP_F1_1; UNKNOWN_1.
 DR PROSITE; PS50262; G_PROTEIN_RECP_F1_2; 1.
 KM Receptor.
 SQ SEQUENCE 331 AA; 37005 MW; 74BA8A8C57859AC CRC64;

Query Match 44.1%; Score 778; DB 2; Length 331;
 Best Local Similarity 53.3%; Pred. No. 4,5e-49;
 Matches 176; Conservative 33; Mismatches 85; Indels 36; Gaps 8;

QY 16 MDSTIPVLGTELTPINGREPTCYKQTLSTFTGLCTIVS-LVALTGNAAVLTMLGCRMRN 74
 DB 28 MDKTIPT-----GSFNGRTLLP-----NLIIIGLVGLTGNAAVFWLLGFRLLARN 72
 QY 75 AVSIYIINLVAADEFLFSGHIIICSPRLNIRHP---ISKILSPWTPPYFGLSMLSAI 131
 DB 73 AFSVYIINLVAADEFLFCHIIIDSTLLKRSYNIIFLPCEFTVMWVPYIAGLSMLSAI 132
 QY 132 STERCLSILMPDIWYHCRPRYLLSSVMCVLWALSLRSILSEMFCDLFSGADSVW---- 187
 DB 133 STERCLSVVCPIWYRCRPRKHTSTVMCSAIVLSLICILNRYFCGL---DTYRKDN 188
 QY 188 -CETSDFTITAMLVFLCVLTCGSSLVLLVRLICGRKMPRLTYVITLTVLVFLCGLP 246
 DB 189 RCLASNFPTAACLIFFLVVLCISLALIVRSFCAGGRKRLRYATIMLVVFLCGLP 248
 QY 247 FGIQWALFSRIHLDKVLFCVHVLVSIFLSALNSANDPIYFVGSFRORONRQNLKVL 306
 DB 249 FGIHWFLLIWKIDYGRKAYGLYLAALVLTAVNSCANPIYFVGSFR-HQKQTLKVL 307
 QY 307 QRALQDTPETAE-----NTVEMSSSKVE 336
 DB 308 QRALQDTPETAE-----NTVEMSSSKVE 330

RESULT 15
 QY 08CDY4 PRELIMINARY; PRT; 338 AA.
 AC 08CDY4;
 DT 01-MAR-2003 (Tremblrel. 23; Last sequence update)
 DT 01-MAR-2003 (Tremblrel. 23; Last sequence update)
 DT 01-MAR-2004 (Tremblrel. 26; Last annotation update)
 DE Mus musculus 0 day neonate head cDNA, RIKEN full-length enriched
 DE library, clone:4833406120 product:MrgB2 G protein-coupled receptor,
 DE full insert sequence.
 GN Name=4833406120Rik;
 OS Mus musculus (Mouse).
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 CC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
 OX NCBI_TaxID=10090;
 RN [1]
 RP SEQUENCE FROM N.A.
 RC STRAIN=C57BL/6J; TISSUE=Head;

This Page Blank (uspio)

GenCore version 5.1.6
Copyright (c) 1993 - 2004 Comphen Ltd.

OM protein - protein search, using sw model

Run on: October 27, 2004, 08:34:24 : Search time 155 Seconds
(without alignments)
779.947 Million cell updates/sec

Title: US-09-867-570-2

Sequence: 1 MESKSWVIRLGLFSLMDSTI.....EGGWLPGTLRLSGRLAQ 337

Scoring table: OLIGO

Gapop 60.0 , Gapext 60.0

Searched: 2002273 seqs, 358729299 residues

Word size : 0

Total number of hits satisfying chosen parameters: 2002273

Minimum DB seq length: 0
Maximum DB seq length: 200000000

Post-processing: Listing first 45 summaries

Database : A_Geneseq_23Sep04:*

1: geneseqp1980s:*\n2: geneseqp1990s:*\n3: geneseqp2000s:*\n4: geneseqp2001s:*\n5: geneseqp2002s:*\n6: geneseqp2003as:*\n7: geneseqp2003bs:*\n8: geneseqp2004s:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	337	100.0	337	5	AAU97598 Human G-P
2	322	95.5	322	3	AAV90761 Human G P
3	322	95.5	322	5	ABU04077 Human G P
4	322	95.5	322	5	AAE21296 Human Mrg
5	322	95.5	322	5	AD116993 Human NOV
6	322	95.5	322	5	AD116998 Human NOV
7	322	95.5	322	6	ADP81750 Human G P
8	322	95.5	322	8	ADH08535 MrGX_3/
9	322	95.5	322	8	AD044602 Human HIT
10	322	95.5	322	8	AD029705 Human GPC
11	322	95.5	560	7	ADP70481 Orphan re
12	277	82.2	322	8	ADF29105 Human GPC
13	274	66.5	322	3	AAE14846 Human nov
14	221	65.6	322	5	AAV90762 Human G P
15	221	65.6	322	5	AD116994 Human NOV
16	221	65.6	322	5	AD116999 Human NOV
17	221	65.6	322	8	AD044604 Human GPC
18	221	65.6	322	8	AD044604 Human HIT
19	181	53.7	302	6	ABP96695 Human G P
20	166	49.3	304	2	AD12766 Human GPC
21	149	44.2	322	2	AAV30159 Human dor
22	100	29.7	318	4	AD12724 Human hum
23	89	26.4	440	4	ABG21073 Novel hum
24	86	25.5	322	2	AAV30160 Human dor
25	76	22.6	76	7	ADD18095 Human G-P

26	76	22.6	328	7	ADD18024 Human G-P
27	50	14.8	319	5	AD116638 Human NOV
28	50	14.8	319	8	ADN42292 Human nov
29	50	14.8	322	2	AAV30162 Human dor
30	50	14.8	322	4	AAE42294 Human GTP
31	50	14.8	322	4	AAE12794 Human G P
32	50	14.8	322	5	AAE17074 Human G-P
33	50	14.8	322	5	AAE17074 Human G-P
34	50	14.8	322	5	ABP95617 Human GPC
35	50	14.8	322	5	AAE21288 Human MrG
36	50	14.8	322	5	AD116991 Human NOV
37	50	14.8	322	5	AD116997 Human NOV
38	50	14.8	322	5	AD116997 Human NOV
39	50	14.8	322	6	ABP96696 Human G P
40	50	14.8	322	6	ABP95266 Human Dor
41	50	14.8	322	7	AD086445 Human GPC
42	50	14.8	322	7	ABW00803 Human GPC
43	50	14.8	322	7	AD196466 Human G P
44	50	14.8	322	8	ADH08520 MrGX1_3/
45	50	14.8	322	8	ADN42290 Human nov

ALIGNMENTS

RESULT 1					
AAU97598 standard; protein; 337 AA.					
XX	AAU97598;				
DT	12-AUG-2002 (first entry)				
XX					
DE	Human G-protein coupled receptor (GPCR).				
XX					
KW	Human, G-protein coupled; receptor; GPCR; human protease;				
KW	human therapeutic protein; query sequence; search; chromosome 3;				
KW	sequence database; non-human transgenic animal; gene therapy.				
XX					
OS	Homo sapiens.				
XX					
PH	Key	Location/Qualifiers			
FT	Domain	41..61	/label= Helix_1		
FT	Domain	75..95	/label= Helix_2		
FT	Domain	112..132	/label= Helix_3		
FT	Modified-site	133..135	/note= "Protein kinase C (PKC) phosphorylation site"		
FT	Domain	151..171	/label= Helix_4		
FT	Modified-site	169..172	/note= "Casein kinase II (CK2) phosphorylation site"		
FT	Modified-site	181..184	/note= "Casein kinase II (CK2) phosphorylation site"		
FT	Domain	193..213	/label= Helix_5		
FT	Modified-site	221..223	/note= "Protein kinase C (PKC) phosphorylation site"		
FT	Domain	229..249	/label= Helix_6		
FT	Modified-site	244..249	/note= "Myristoylation site"		
FT	Modified-site	248..253	/note= "Myristoylation site"		
FT	Domain	261..281	/label= Helix_7		
FT	Modified-site	279..282	/note= "Ser phosphorylation site"		
FT	Modified-site	292..294	/note= "Protein kinase C (PKC) phosphorylation site"		
FT	Modified-site	333..336	/note= "Protein kinase C (PKC) phosphorylation site"		

/note= "Casein kinase II (CK2) phosphorylation site"

MO200234914-A1.

02-MAY-2002.

10-OCT-2001; 2001WO-US031592.

25-OCT-2000; 2000US-00695045.
31-MAY-2001; 2001US-00867570.

(PEKE) PE CORP NY.

Wei M, Zhao Q, Woodage T, Di Francesco V, Beasley EM;

WPI; 2002-463360/49.

N-PSDB; ABK52822, ABK52823.

Novel isolated G-protein coupled receptor peptide useful for treating disorder characterized by absence of, in appropriate or unwanted expression of the receptor protein, and as immunogens to raise antibodies.

Claim 1; Fig 2; 75bp; English.

The present invention relates to a new G-protein coupled receptor (GPCR) peptide. The invention is useful for identifying a modulator of GPCR and for treating a disease or condition mediated by a human protease. The invention is also useful as models for the development of human therapeutics, for identifying therapeutic proteins, as targets for development of human therapeutic agents, and as query sequence to perform a search against sequence databases to, for e.g., identify other family members of related sequences. The vector of the invention is useful for producing a GPCR protein or peptide, for conducting cell-based assays involving the GPCR protein or its fragment, for identifying GPCR protein mutants whose functions are affected, and to produce non-human transgenic animals. The present amino acid sequence represents the human G-protein coupled receptor (GPCR) protein of the invention. This sequence is encoded by the human G-protein coupled receptor (GPCR) gene located on chromosome 3

Sequence 337 AA;

Query Match 100.0%; Score 337; DB 5; Length 337;

Best Local Similarity 100.0%; Pred. No. 0; Matches 337; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

```

QY 1 MESKSSWIRIGFPLSMDSITPIVGTETPPINGREBTPCKQTLSTFGLTCIVSLVATGN 60
DB 1 MESKSSWIRIGFPLSMDSITPIVGTETPPINGREBTPCKQTLSTFGLTCIVSLVATGN 60
QY 61 AVVAMLGCRMRNRNAVSIYILNVAADFLFSGHIICSPRLINIRHPISTKILSPVMPFP 120
DB 61 AVVAMLGCRMRNRNAVSIYILNVAADFLFSGHIICSPRLINIRHPISTKILSPVMPFP 120
QY 121 YFGLSMLSAISTERCLSIWPVWYHCRPRYLSVWCYLWALSILRSILEMFCDFLF 180
DB 121 YFGLSMLSAISTERCLSIWPVWYHCRPRYLSVWCYLWALSILRSILEMFCDFLF 180
QY 181 SGADSWCETSDFTITAMLVFLCSSLVLVRLICGSRKMPETRLRYTITLLTVLVF 240
DB 181 SGADSWCETSDFTITAMLVFLCSSLVLVRLICGSRKMPETRLRYTITLLTVLVF 240
QY 241 LLGGLPFGIOALFRIHLDWKFVFCVHLVSIIFLSALNSSANPIIYFVGSFRORONQ 300
DB 241 LLGGLPFGIOALFRIHLDWKFVFCVHLVSIIFLSALNSSANPIIYFVGSFRORONQ 300
QY 301 NLKVLQRLADTPEVDEGGGWLPGETLELSSRLLEQ 337
DB 301 NLKVLQRLADTPEVDEGGGWLPGETLELSSRLLEQ 337

```

RESULT 2

AA90761 ID AA90761 standard; protein, 322 AA.

AA90761;

18-AUG-2000 (first entry)

Human G protein-coupled receptor hH17T213 SEQ ID NO:1.

Human; G protein-coupled receptor; hippocampus; diagnosis; screening; genetic disease; cellular function regulation.

Homo sapiens.

WO200020455-A1.

13-APR-2000.

30-SEP-1999; 99WO-JP005366.

01-OCT-1998; 98JP-00279535.

(TAKE) TAKEDA CHEM IND LTD.

Watanabe T, Terao Y, Matsui H;

WPI; 2000-303747/26.

N-PSDB; AAA29811.

Human-derived G protein-coupled protein and encoding nucleic acid, useful e.g. in determining ligands and treatment of diseases associated with dysfunction of the protein.

Claim 1; Page 90-91; 97bp; Japanese.

The present sequence represents a human-derived G protein-coupled protein designated hH17T213, which is isolated from the human hippocampus. The G protein-coupled receptor can be used for preventing, treating and diagnosing genetic diseases associated with G protein-coupled protein, and for regulating cellular functions. The protein can be used to prevent and treat disorders associated with G protein-coupled protein gene dysfunction. It can also be used to identify G protein-coupled protein ligands and generating antibodies and antisera against the protein. It is also useful in constructing recombinant receptor protein expression systems, developing receptor-binding assay systems and screening drug candidates, and can be used as a probe in the genetic diagnosis of G protein-coupled protein disorders

Sequence 322 AA;

Query Match 95.5%; Score 322; DB 3; Length 322;

Best Local Similarity 100.0%; Pred. No. 2,9e-309; Matches 322; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

```

QY 16 MDSTIPVIGTETPPINGREBTPCKQTLSTFGLTCIVSLVATGNNAVVLGCRMRNRNA 75
DB 1 MDSTIPVIGTETPPINGREBTPCKQTLSTFGLTCIVSLVATGNNAVVLGCRMRNRNA 75
QY 76 VSIYILNVAADFLFSGHIICSPRLINIRHPISTKILSPVMPFPYFGLSMLSAISTER 135
DB 76 VSIYILNVAADFLFSGHIICSPRLINIRHPISTKILSPVMPFPYFGLSMLSAISTER 135
QY 136 CLSILMPVWYHCRPRYLSVWCYLWALSILRSILEMFCDFLFGADSWCETSDFTIT 195
DB 136 CLSILMPVWYHCRPRYLSVWCYLWALSILRSILEMFCDFLFGADSWCETSDFTIT 195
QY 195 61 VSIYILNVAADFLFSGHIICSPRLINIRHPISTKILSPVMPFPYFGLSMLSAISTER 240
DB 195 61 VSIYILNVAADFLFSGHIICSPRLINIRHPISTKILSPVMPFPYFGLSMLSAISTER 240
QY 240 181 IAMLVFLCSSLVLVRLICGSRKMPETRLRYTITLLTVLVFLCGPFGIOALF 255
DB 240 181 IAMLVFLCSSLVLVRLICGSRKMPETRLRYTITLLTVLVFLCGPFGIOALF 255
QY 255 256 RIHLDWKFVFCVHLVSIIFLSALNSSANPIIYFVGSFRORONKLVQLQRLADTPE 315
DB 256 241 RIHLDWKFVFCVHLVSIIFLSALNSSANPIIYFVGSFRORONKLVQLQRLADTPE 300

```

QY 316 VDEGGMLPQETLELSGSRLEQ 337
DB 301 VDEGGMLPQETLELSGSRLEQ 322

RESULT 3
ABJ04077
ID ABJ04077 standard; protein; 322 AA.
AC ABJ04077;
DT 11-OCT-2002 (first entry)
DE Human G protein coupled receptor hrUP37.
XX
XX
XX Human; G-protein coupled receptor; GPCR; hrUP28; hrUP29; hrUP30; hrUP31;
KM hrUP32; hrUP33; hrUP34; hrUP35; hrUP36; hrUP37.
XX
XX Homo sapiens.
OS
PN WO200242461-A2.
XX
XX 30-MAY-2002.
XX
XX 26-NOV-2001; 2001WO-US044386.
XX
XX 27-NOV-2000; 2000US-0253404P.
PR 12-DEC-2000; 2000US-0253366P.
PR 20-FEB-2001; 2001US-0270266P.
PR 20-FEB-2001; 2001US-0270286P.
PR 06-APR-2001; 2001US-0282032P.
PR 06-APR-2001; 2001US-0282356P.
PR 06-APR-2001; 2001US-0282358P.
PR 06-APR-2001; 2001US-0282365P.
PR 14-MAY-2001; 2001US-0290917P.
PR 31-JUL-2001; 2001US-0309208P.
XX
XX (AREN-) ARENA PHARM INC.
PA
PI Chen R, Chu ZL, Dang HT, Lowitz KP, Pride C;
XX
XX MPI; 2002-566565/60.
DR N-PSDB; ABT04875.
XX
XX Novel endogenous and non-endogenous versions of G protein-coupled
PT receptor useful for identification of candidate compounds as receptor
PT agonists or antagonists for use as therapeutic agents.
XX
XX Claim 37, Page 75-76; 84pp; English.
PS
XX The present invention provides the protein and coding sequences of
CC several human G-protein coupled receptors (GPCRs). These can be used in
CC the identification of candidate compounds as receptor agonists or inverse
CC agonists having applicability as therapeutic agents. The present sequence
CC is a GPCR protein of the invention
XX
XX Sequence 322 AA;
SQ

Query Match 95.5%; Score 322; DB 5; Length 322;
Best Local Similarity 100.0%; Pred. No. 2.9e-309;
Matches 322; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 16 MDSTIPVLGTELPINGRETPCYKOTLSFTGTCIVSLVATGNNAVVTMLGCRMRNA 75
DB 1 MDSTIPVLGTELPINGRETPCYKOTLSFTGTCIVSLVATGNNAVVTMLGCRMRNA 60
QY 76 VSIYIINLVADFLFSGHIIICSPILRINIRHPISTKILSPVMTFPYRIGLSMISAISTER 135
DB 61 VSIYIINLVADFLFSGHIIICSPILRINIRHPISTKILSPVMTFPYRIGLSMISAISTER 120
QY 136 CUSILPWIWHCRPRYLSVMCVLLMALSLRSILEMFCDFLFGADSVWCETSDPIT 195

DB 121 CUSILPWIWHCRPRYLSVMCVLLMALSLRSILEMFCDFLFGADSVWCETSDPIT 180
QY 196 IAMLVFLCVLLCGSSLVLRILCGSRKMPRLRYTITLLTVLVLGCLPFGIOWALFS 255
DB 181 IAMLVFLCVLLCGSSLVLRILCGSRKMPRLRYTITLLTVLVLGCLPFGIOWALFS 240
QY 256 RIHLDMKVLFCFHVHLYSIFLSALNSSANPIIYFVGSFPRORONRQNLQBALQDTPR 315
DB 241 RIHLDMKVLFCFHVHLYSIFLSALNSSANPIIYFVGSFPRORONRQNLQBALQDTPR 300
QY 316 VDEGGMLPQETLELSGSRLEQ 337
DB 301 VDEGGMLPQETLELSGSRLEQ 322

RESULT 4
AAE21296
ID AAE21296 standard; protein; 322 AA.
XX
XX AAE21296;
AC
XX
XX 01-JUL-2002 (first entry)
DT
XX
XX Human MrgX3 (mas-related gene) protein.
DE
XX
XX Human; mas-related gene; G-protein coupled receptor; drg-12 protein;
KM receptor; sensory perception; pain; analgesic; MrgX3.
XX
XX Homo sapiens.
OS
PN WO200183555-A2.
XX
XX 08-NOV-2001.
XX
XX 04-MAY-2001; 2001WO-US014519.
XX
XX 04-MAY-2000; 2000US-0202027P.
PR 01-AUG-2000; 2000US-0222344P.
PR 03-NOV-2000; 2000US-00704707.
PR 19-APR-2001; 2001US-0285493P.
XX
XX (CALY) CALIFORNIA INST OF TECHNOLOGY.
PA
PI Anderson DJ, Dong X, Zylka M, Han S, Simon M;
XX
XX MPI; 2002-171346/22.
DR N-PSDB; AAD33751.
XX
XX Isolated polypeptide, Mrg, which is a G-protein coupled receptor and an
PT isolated polypeptide, drg-12, which is also a receptor, useful for
PT identifying agonists or antagonists for treating pain.
XX
XX Claim 16; Page 130; 185pp; English.
PS
XX The invention relates to Mrg (mas-related gene) protein, which is a G-
CC protein coupled receptor and drg-12 protein, which is a receptor. The
CC invention is useful for identifying compounds that bind to it, especially
CC agonists or antagonists. Administration of an agent (e.g. the identified
CC agonist) that increases the expression of Mrg in a mammal may be used for
CC treating impaired sensory perception in a mammal, especially pain. The
CC antagonist may also be useful for treating impaired sensory perception in
CC a mammal. The present sequence is human MrgX3 protein
XX
XX Sequence 322 AA;
SQ

Query Match 95.5%; Score 322; DB 5; Length 322;
Best Local Similarity 100.0%; Pred. No. 2.9e-309;
Matches 322; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 16 MDSTIPVLGTELPINGRETPCYKOTLSFTGTCIVSLVATGNNAVVTMLGCRMRNA 75
DB 1 MDSTIPVLGTELPINGRETPCYKOTLSFTGTCIVSLVATGNNAVVTMLGCRMRNA 60

QY 76 VSIYIILNVAADFLFLSGHIICSPRLINIRHPISKILSPVTPFYFGLSMLSAISTER 135
 DB 61 VSIYIILNVAADFLFLSGHIICSPRLINIRHPISKILSPVTPFYFGLSMLSAISTER 120
 QY 136 CLSTLMPWYHCRPRVYSSWCVLLMLSLRSLLEWPCDFLPSGADSWCERSDFTT 195
 DB 121 CLSTLMPWYHCRPRVYSSWCVLLMLSLRSLLEWPCDFLPSGADSWCERSDFTT 180
 QY 196 IAWLVFLCVVCGSSLVLLVRLICGSRKMPFLRLVYVTLILFVLVFLCGLPFGIOMALFS 255
 DB 181 IAWLVFLCVVCGSSLVLLVRLICGSRKMPFLRLVYVTLILFVLVFLCGLPFGIOMALFS 240
 QY 256 RIHLDMKVLFCGHVHVSIFLSALNSSANPIYFVGSFRORONRKLVLQRALQDTP 315
 DB 241 RIHLDMKVLFCGHVHVSIFLSALNSSANPIYFVGSFRORONRKLVLQRALQDTP 300
 QY 316 VDEGGGWLPORTLISGSRLEQ 337
 DB 301 VDEGGGWLPORTLISGSRLEQ 322

RESULT 5
 AD116993
 ID AD116993 standard; protein, 322 AA.
 AC AD116993;
 XX
 DT 15-APR-2004 (first entry)
 XX
 DE Human NOVA protein homologue SegID 529.
 XX
 KW human; NOVA; cardiomyopathy; atherosclerosis; cancer; diabetes;
 KW inflammation; autoimmune disorder; allergy; blood disorder;
 KW acquired immunodeficiency syndrome; AIDS; obesity; asthma;
 KW immunoglobulin (Ig)A nephropathy; cirrhosis; arthritis;
 KW Alzheimer's disease; infection; str.
 XX
 OS Homo sapiens.
 XX
 PN MO200268649-A2.
 XX
 PD 06-SEP-2002.
 XX
 PF 31-JAN-2002; 2002WO-US002785.
 XX
 PR 31-JAN-2001; 2001US-0265395P.
 PR 31-JAN-2001; 2001US-0265412P.
 PR 31-JAN-2001; 2001US-0265514P.
 PR 31-JAN-2001; 2001US-0265517P.
 PR 02-FEB-2001; 2001US-0266406P.
 PR 05-FEB-2001; 2001US-0266767P.
 PR 07-FEB-2001; 2001US-0266975P.
 PR 07-FEB-2001; 2001US-0267057P.
 PR 08-FEB-2001; 2001US-0267459P.
 PR 09-FEB-2001; 2001US-0267823P.
 PR 15-FEB-2001; 2001US-0268974P.
 PR 26-FEB-2001; 2001US-0271664P.
 PR 27-FEB-2001; 2001US-0271839P.
 PR 27-FEB-2001; 2001US-0271855P.
 PR 02-MAR-2001; 2001US-0272788P.
 PR 02-MAR-2001; 2001US-0273046P.
 PR 14-MAR-2001; 2001US-0275925P.
 PR 14-MAR-2001; 2001US-0275947P.
 PR 14-MAR-2001; 2001US-0275950P.
 PR 14-MAR-2001; 2001US-0275989P.
 PR 15-MAR-2001; 2001US-0276448P.
 PR 15-MAR-2001; 2001US-0276450P.
 PR 16-MAR-2001; 2001US-0276397P.
 PR 16-MAR-2001; 2001US-0276768P.
 PR 20-MAR-2001; 2001US-0278652P.
 PR 26-MAR-2001; 2001US-0278775P.
 PR 29-MAR-2001; 2001US-0279882P.

PR 29-MAR-2001; 2001US-0279884P.
 PR 30-MAR-2001; 2001US-0280147P.
 PR 11-APR-2001; 2001US-0282929P.
 PR 11-APR-2001; 2001US-0283083P.
 PR 20-APR-2001; 2001US-0285133P.
 PR 23-APR-2001; 2001US-0285749P.
 PR 03-MAY-2001; 2001US-0288327P.
 PR 03-MAY-2001; 2001US-0288504P.
 PR 29-MAY-2001; 2001US-0294047P.
 PR 30-MAY-2001; 2001US-0294473P.
 PR 08-JUN-2001; 2001US-0296964P.
 PR 18-JUN-2001; 2001US-0298959P.
 PR 19-JUN-2001; 2001US-0299324P.
 PR 13-AUG-2001; 2001US-0312020P.
 PR 16-AUG-2001; 2001US-0312889P.
 PR 16-AUG-2001; 2001US-0312908P.
 PR 21-AUG-2001; 2001US-0313390P.
 PR 28-AUG-2001; 2001US-0315470P.
 PR 31-AUG-2001; 2001US-0316447P.
 PR 07-SEP-2001; 2001US-0318115P.
 PR 07-SEP-2001; 2001US-0318118P.
 PR 12-SEP-2001; 2001US-0318740P.
 PR 19-SEP-2001; 2001US-0323379P.
 PR 18-OCT-2001; 2001US-0330245P.
 PR 18-OCT-2001; 2001US-0330308P.
 PR 14-NOV-2001; 2001US-0332701P.

(CURA-) CURAGEN CORP.
 XX
 PA
 XX
 PI Tchermey VT, Szytek KA, Zernusen BD, Patturajan M, Shinkets RA;
 PI Li U, Gangoli EA, Padigaru M, Anderson DM, Rastelli L, Miller CB;
 PI Gerlach VL, Taupier RV, Gusev VY, Colman SD, Wolenc AR, Pena CE;
 PI Furtak K, Grose WM, Albrook JP, Lepley DM, Rieger DK, Burgess CE;
 XX
 DR MPI; 2002-706998/76.
 XX
 PT New NOVA polypeptides and nucleic acids, useful for preventing or
 PT treating NOVA-associated disorders, e.g. cancer, cardiomyopathy,
 PT atherosclerosis, or diabetes, and in chromosome mapping, tissue typing or
 PT pharmacogenomics.
 XX
 PS Disclosure; SEQ ID NO 529; 1498bp; English.

This invention relates to a novel nucleic acids, and encoded polypeptides thereof, which have properties related to the stimulation of biochemical or physiological responses in a cell, tissue, organ or organism. Specifically, it refers to the use of biologically active fragments for diagnostic and prognostic assays and furthermore in the treatment of diverse pathological conditions. The present invention describes novel human and murine NOVA proteins, as well as methods to modulate their expression using antisense oligos, ribozymes and peptide nucleic acids. The NOVA polypeptides, polynucleotides and antibodies are useful in treating or preventing NOVA-associated disorders, e.g. cardiomyopathy, atherosclerosis, cancer and diabetes. Furthermore, they may be used in treating or preventing diseases such as inflammation, autoimmune disorders, allergies, blood disorders, acquired immunodeficiency syndrome (AIDS), obesity, aschma, immunoglobulin (Ig)A nephropathy, cirrhosis, arthritis, Alzheimer's disease, infections, stroke, muscular dystrophy and epilepsy. Accordingly, these molecules have many activities including cytostatic, cardiac, anti-inflammatory, immunosuppressive, antiallergic, haemostatic, anti-HIV, antidiabetic, antiatherosclerotic, anorectic, antiaesthetic, nephrotoxic, antibacterial, virucide, antiparasitic, neuroprotective, nootropic, anticholinergic, anticholinergic, anticholinergic, relaxant and anticonvulsant. In addition, they are useful in screening assays to identify small molecules that modulate or inhibit, for example, neurogenesis, wound healing and angiogenesis. The nucleic acids are also used as in chromosome mapping, tissue typing, preventive medicine and pharmacogenomics. This polypeptide is a homologue of a human NOVA protein of the invention.

Query Match 95.5%; Score 322; DB 5; Length 322;

Best Local Similarity 100.0%; Pred. No. 2.9e-309;
Matches 322; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 16 MDSTIPVGLGRLPPINGRETPCKQKLTSTGLCTIVSLVAGNAVVMILGCRMRNA 75
DB 1 MDSTIPVGLGRLPPINGRETPCKQKLTSTGLCTIVSLVAGNAVVMILGCRMRNA 60
QY 76 VSIYIINLVADFLPSGHIICSPRLINIRHPISKLSIPVMTPEYPIGSMLSAISTER 135
DB 61 VSIYIINLVADFLPSGHIICSPRLINIRHPISKLSIPVMTPEYPIGSMLSAISTER 120
QY 136 CLSLIPWVHCRPRPRLTSSVWCYLWALSILRSIIEWMFCDFLPSGADSVWCETSDFIT 195
DB 121 CLSLIPWVHCRPRPRLTSSVWCYLWALSILRSIIEWMFCDFLPSGADSVWCETSDFIT 180
QY 196 IANLVPLCVLLCGSSLVLLVRIICGSRKMP/LRYLYTITLVVPLLCG/PFGIQWALFS 255
DB 181 IANLVPLCVLLCGSSLVLLVRIICGSRKMP/LRYLYTITLVVPLLCG/PFGIQWALFS 240
QY 256 RIHLDMKVLFCCHVHLSVIFLSALNSSANPITTYFVGSFRORONRNLK/LORALODTPE 315
DB 241 RIHLDMKVLFCCHVHLSVIFLSALNSSANPITTYFVGSFRORONRNLK/LORALODTPE 300
QY 316 VDEGGGWLPOETLELSSGRLEQ 337
DB 301 VDEGGGWLPOETLELSSGRLEQ 322

RESULT 6
AD116998
ID AD116998 standard; protein; 322 AA.
AC AD116998;
XX
DT 15-APR-2004 (first entry)
XX
DE Human NOVX protein homologue Segid 534.
XX
KW human; NOVX; cardiomyopathy; atherosclerosis; cancer; diabetes;
KW inflammation; autoimmune-disorder; allergy; blood disorder;
KW acquired immunodeficiency syndrome; AIDS; obesity; asthma;
KW immunoglobulin (Ig)A nephropathy; cirrhosis; arthritis;
KW Alzheimer's disease; infection; str.
XX
OS Homo sapiens.
XX
PN WO200268649-A2.
XX
PD 06-SEP-2002.
XX
PF 31-JAN-2002; 2002MO-US002785.
XX
PR 31-JAN-2001; 2001US-0265395P.
XX
PR 31-JAN-2001; 2001US-0265412P.
XX
PR 31-JAN-2001; 2001US-0265514P.
XX
PR 31-JAN-2001; 2001US-0265517P.
XX
PR 02-FEB-2001; 2001US-0266406P.
XX
PR 05-FEB-2001; 2001US-0266767P.
XX
PR 07-FEB-2001; 2001US-0267057P.
XX
PR 08-FEB-2001; 2001US-0267459P.
XX
PR 09-FEB-2001; 2001US-0267823P.
XX
PR 15-FEB-2001; 2001US-0268974P.
XX
PR 26-FEB-2001; 2001US-0271664P.
XX
PR 27-FEB-2001; 2001US-0271839P.
XX
PR 27-FEB-2001; 2001US-0271855P.
XX
PR 02-MAR-2001; 2001US-0272788P.
XX
PR 02-MAR-2001; 2001US-0273046P.
XX
PR 14-MAR-2001; 2001US-0275925P.
XX
PR 14-MAR-2001; 2001US-0275947P.
XX
PR 14-MAR-2001; 2001US-0275950P.
XX
PR 14-MAR-2001; 2001US-0275989P.
XX
PR 15-MAR-2001; 2001US-0276448P.

PR 15-MAR-2001; 2001US-0276450P.
PR 16-MAR-2001; 2001US-0276397P.
PR 16-MAR-2001; 2001US-0276768P.
PR 20-MAR-2001; 2001US-0278652P.
PR 26-MAR-2001; 2001US-0278775P.
PR 26-MAR-2001; 2001US-0278778P.
PR 29-MAR-2001; 2001US-0279682P.
PR 29-MAR-2001; 2001US-0279684P.
PR 30-MAR-2001; 2001US-0280147P.
PR 11-APR-2001; 2001US-0282992P.
PR 11-APR-2001; 2001US-0283083P.
PR 20-APR-2001; 2001US-0285133P.
PR 23-APR-2001; 2001US-0285749P.
PR 03-MAY-2001; 2001US-0288327P.
PR 03-MAY-2001; 2001US-0288504P.
PR 29-MAY-2001; 2001US-0294047P.
PR 30-MAY-2001; 2001US-0294473P.
PR 08-JUN-2001; 2001US-0296964P.
PR 18-JUN-2001; 2001US-0298959P.
PR 19-JUN-2001; 2001US-0299324P.
PR 13-AUG-2001; 2001US-0312020P.
PR 16-AUG-2001; 2001US-0312889P.
PR 16-AUG-2001; 2001US-0312908P.
PR 21-AUG-2001; 2001US-0313390P.
PR 28-AUG-2001; 2001US-0315470P.
PR 31-AUG-2001; 2001US-0316447P.
PR 07-SEP-2001; 2001US-0318115P.
PR 07-SEP-2001; 2001US-0318118P.
PR 12-SEP-2001; 2001US-0318740P.
PR 19-SEP-2001; 2001US-0323379P.
PR 18-OCT-2001; 2001US-0330245P.
PR 18-OCT-2001; 2001US-0330308P.
PR 14-NOV-2001; 2001US-0332701P.

(CURA-) CURAGEN CORP.
XX
PA Tchernev VT, Spytek KA, Zernhusen BD, Patumrajan M, Shinkets RA;
XX
PI Li L, Gangolli EA, Padgugan M, Anderson DW, Raetelli L, Miller CE;
XX
PI Gerlach VL, Taulier RJ, Gusev VY, Colman SD, Wolenc AR, Pena CE;
XX
PI Firtak K, Grosse WM, Alsbrook UP, Lepley DM, Rieger DK, Burgess CE;
XX
DR WPI; 2002-706998/76.
XX
XX New NOVX polypeptides and nucleic acids, useful for preventing or
PT treating NOVX-associated disorders, e.g. cancer, cardiomyopathy,
PT atherosclerosis, or diabetes, and in chromosome mapping, tissue typing or
PT pharmacogenomics.
XX
PS Disclosure; SEQ ID NO 534; 1498bp; English.
XX
XX This invention relates to a novel nucleic acids, and encoded polypeptides
CC thereof, which have properties related to the stimulation of biochemical
CC or physiological responses in a cell, tissue, organ or organism.
CC Specifically, it refers to the use of biologically active fragments for
CC diagnostic and prognostic assays and furthermore in the treatment of
CC diverse pathological conditions. The present invention describes novel
CC human and murine NOVX proteins, as well as methods to modulate their
CC expression using antisense oligos, ribozymes and peptide nucleic acids.
CC The NOVX polypeptides, polynucleotides and antibodies are useful in
CC treating or preventing NOVX-associated disorders, e.g. cardiomyopathy,
CC atherosclerosis, cancer and diabetes. Furthermore, they may be used in
CC treating or preventing diseases such as inflammation, autoimmune
CC disorders, allergies, blood disorders, acquired immunodeficiency syndrome
CC (AIDS), obesity, asthma, immunoglobulin (Ig)A nephropathy, cirrhosis,
CC arthritis, Alzheimer's disease, infections, stroke, muscular dystrophy
CC and epilepsy. Accordingly, these molecules have many activities including
CC cytostatic, cardiant, antiinflammatory, immunosuppressive, antiallergic,
CC hemostatic, anti-HIV, antidiabetic, antiarteriosclerotic, anorectic,
CC antistimatic, nephrotropic, antiarthritic, hepatotropic,
CC neuroprotective, nootropic, antibacterial, virucide, antiparasitic,
CC relaxant and anticonvulsant. In addition, they are useful in screening
CC assays to identify small molecules that modulate or inhibit, for example,
CC neurogenesis, wound healing and angiogenesis. The nucleic acids are also

CC used as in chromosome mapping, tissue typing, preventive medicine and
CC pharmacogenomics. This polypeptide is a homologue of a human NOXA protein
CC of the invention.
XX
XX

Sequence 322 AA:

Query Match 95.5%; Score 322; DB 5; Length 322;
Best Local Similarity 100.0%; Pred. No. 2.9e-309; Indels 0; Gaps 0;
Matches 322; Conservative 0; Mismatches 0;

QY 16 MDSTIPVAGTELTPTNGREETPCYKQTLSFTGLTCIVSLVATLGNAAVLMILGCRMRNA 75
Db 1 MDSTIPVAGTELTPTNGREETPCYKQTLSFTGLTCIVSLVATLGNAAVLMILGCRMRNA 60
QY 76 VSIYILNVADFLPLSGHIIICSPRLINIRHPISKILSPVMTFPYFGLSMLSAISTER 135
Db 61 VSIYILNVADFLPLSGHIIICSPRLINIRHPISKILSPVMTFPYFGLSMLSAISTER 120
QY 136 CLSIIMPWTWYHCRPRRYLSSVWCVLWALSLRSIILEMFCDFLPSGADSVWCETSDFT 195
Db 121 CLSIIMPWTWYHCRPRRYLSSVWCVLWALSLRSIILEMFCDFLPSGADSVWCETSDFT 180
QY 196 IAMLVFLCVVLCSSSLVLLVRLICGSRKMPRLRYVTILTLVFLVFLGCLPFGIOMALFS 255
Db 181 IAMLVFLCVVLCSSSLVLLVRLICGSRKMPRLRYVTILTLVFLVFLGCLPFGIOMALFS 240
QY 256 RIHLDWKVLFCHVHVSIFLSALNSSANPIIYFVGSFRORONKLVLRALQDTPPE 315
Db 241 RIHLDWKVLFCHVHVSIFLSALNSSANPIIYFVGSFRORONKLVLRALQDTPPE 300
QY 316 VDEGGWLPQETLELSGSRLEQ 337
Db 301 VDEGGWLPQETLELSGSRLEQ 322

RESULT 7

ABP81750
ID ABP81750 standard; protein; 322 AA.

AC ABP81750;

DT 04-MAR-2003 (first entry)

DE Human G protein-coupled receptor MrgX3 protein SEQ ID NO:674.

XX G protein-coupled receptor; GPCR; antigenic peptide; gene therapy;
KM G protein-coupled receptor modulator; antibody; immune-related disease;
KM growth-related disease; cell regeneration-related disease; AIDS; cancer;
KM immunological-related cell proliferative disease; autoimmune disease;
KM Alzheimer's disease; atherosclerosis; infection; osteoarthritis; allergy;
KM osteoporosis; cardiomyopathy; inflammation; Crohn's disease; diabetes;
KM graft versus host disease; Parkinson's disease; multiple sclerosis; pain;
KM psoriasis; anxiety; depression; schizophrenia; dementia; memory loss;
KM mental retardation; epilepsy; asthma; tuberculosis; obesity; nausea;
KM hypertension; hypotension; renal disorder; rheumatoid arthritis; trauma;
ulcer.

XX Homo sapiens.

XX W0200261087-A2.

XX 08-AUG-2002.

XX 19-DEC-2001; 2001WO-US050107.

XX 19-DEC-2000; 2000US-0257144P.

XX (LIFE-) LIFESPAN BIOSCIENCES INC.

XX Burnet GC, Roush CL, Brown JP;

XX WPI: 2003-046718/04.

XX N-PSDB; ABZ42595.

XX New isolated antigenic peptides e.g., for G protein-coupled receptors
PT (GPCR), useful for diagnosing and designing drugs for treating conditions
PT in which GPCRs are involved, e.g. AIDS, Alzheimer's disease, cancer or
PT autoimmune diseases.

PS Disclosure: Fig 1; 523pp; English.

XX The present invention describes antigenic peptides (1) comprising: (a)
CC any one of 1601 sequences (see ABP82019 to ABP83619) of 12-24 amino
CC acids. Also described: (1) an assay for the detection of a particular G
CC protein-coupled receptor (GPCR) or a candidate polypeptide in a sample;
CC and (2) an isolated antibody having high specificity and high affinity or
CC avidity for a particular GPCR. (1) can be used as GPCR modulators and in
CC gene therapy. The antigenic peptides for GPCRs are useful in detecting an
CC antibody against a particular GPCR, and in the production of specific
CC antibodies. The peptides and antibodies are also useful for detecting the
CC presence or absence of corresponding GPCRs. The antigenic peptides for
CC GPCRs and antibodies are useful for diagnosing and designing drugs for
CC treating immune-related diseases, growth-related diseases, cell
CC regeneration-related disease, immunological-related cell proliferative
CC diseases, or autoimmune diseases, e.g. AIDS, Alzheimer's disease,
CC atherosclerosis, bacterial, fungal, protozoan or viral infections,
CC osteoarthritis, osteoporosis, cancer, cardiomyopathy, chronic and acute
CC inflammation, allergies, Crohn's disease, diabetes, grapt versus host
CC disease, Parkinson's disease, multiple sclerosis, pain, psoriasis,
CC anxiety, depression, schizophrenia, dementia, mental retardation, memory
CC loss, epilepsy, asthma, tuberculosis, obesity, nausea, hypertension,
CC hypotension, renal disorders, rheumatoid arthritis, trauma, ulcers, or
CC any other disorder in which GPCRs are involved. The antibodies may be
CC used in immunoassays and immunodiagnosis. ABZ42523 to ABZ42869 encode
CC GPCR proteins given in ABP81675 to ABP82018, which are used in the
CC exemplification of the present invention
XX

Sequence 322 AA:

Query Match 95.5%; Score 322; DB 5; Length 322;
Best Local Similarity 100.0%; Pred. No. 2.9e-309; Indels 0; Gaps 0;
Matches 322; Conservative 0; Mismatches 0;

QY 16 MDSTIPVAGTELTPTNGREETPCYKQTLSFTGLTCIVSLVATLGNAAVLMILGCRMRNA 75
Db 1 MDSTIPVAGTELTPTNGREETPCYKQTLSFTGLTCIVSLVATLGNAAVLMILGCRMRNA 60
QY 76 VSIYILNVADFLPLSGHIIICSPRLINIRHPISKILSPVMTFPYFGLSMLSAISTER 135
Db 61 VSIYILNVADFLPLSGHIIICSPRLINIRHPISKILSPVMTFPYFGLSMLSAISTER 120
QY 136 CLSIIMPWTWYHCRPRRYLSSVWCVLWALSLRSIILEMFCDFLPSGADSVWCETSDFT 195
Db 121 CLSIIMPWTWYHCRPRRYLSSVWCVLWALSLRSIILEMFCDFLPSGADSVWCETSDFT 180
QY 196 IAMLVFLCVVLCSSSLVLLVRLICGSRKMPRLRYVTILTLVFLVFLGCLPFGIOMALFS 255
Db 181 IAMLVFLCVVLCSSSLVLLVRLICGSRKMPRLRYVTILTLVFLVFLGCLPFGIOMALFS 240
QY 256 RIHLDWKVLFCHVHVSIFLSALNSSANPIIYFVGSFRORONKLVLRALQDTPPE 315
Db 241 RIHLDWKVLFCHVHVSIFLSALNSSANPIIYFVGSFRORONKLVLRALQDTPPE 300
QY 316 VDEGGWLPQETLELSGSRLEQ 337
Db 301 VDEGGWLPQETLELSGSRLEQ 322

RESULT 8

ADH08535
ID ADH08535 standard; protein; 322 AA.

AC ADH08535;

DT 25-MAR-2004 (first entry)

DE MrGX3.
XX mas-related gene D; MrGD; Analgesic; Vulnerary; Ophthalmological;
KM sensory perception; glaucoma; MrG.
XX Mus musculus.
XX WO2004003133-A1.
XX 08-JAN-2004.
XX 13-MAY-2003; 2003WO-US015004.
XX 26-JUN-2002; 2002US-00183116.
XX (CALY) CALIFORNIA INST OF TECHNOLOGY.
XX Anderson DJ, Dong X, Zylka M, Han S, Simon MI;
XX WPI; 2004-083025/08.
XX N-PSDB; ADH08534.
XX
XX New mas-related gene D polypeptides, useful as therapeutics or in
PT identifying agonists or antagonists that alter pain perception in a
PT mammal for treating impaired sensory perception, e.g. chronic intractable
PT pain or neuropathic pain.
XX
XX Disclosure; SEQ ID NO 31; 220pp; English.
XX
XX The present invention relates to an isolated mas-related gene D (MrGD)
CC polypeptide. The MrGD polypeptides are useful as therapeutics or for
CC identifying compounds, i.e. agonists or antagonists, that alter pain
CC perception in a mammal. The compounds are useful for treating impaired
CC sensory perception, e.g. chronic intractable pain or neuropathic pain,
CC promoting wound healing, restoring normal sensitivity following injury,
CC or treating ocular conditions, particularly those associated with
CC pressure such as glaucoma. The MrG genes or proteins may be used as
CC molecular probes for the detection of cells or tissues related to or
CC involved with sensory perception. The present sequence represents a MrGA
CC (MrG subfamily) protein.
XX
XX Sequence 322 AA;
SQ
Query Match 95.5%; Score 322; DB 8; Length 322;
Best Local Similarity 100.0%; Pred. No. 2.9e-309; Indels 0; Gaps 0;
Matches 322; Conservative 0; Mismatches 0;
QY 16 MDSTIPVLGTELPINGREBTPCYKOTLSFTGLTCTVSLVALTGNVAVMLLGCRRRNA 75
DB 1 MDSTIPVLGTELPINGREBTPCYKOTLSFTGLTCTVSLVALTGNVAVMLLGCRRRNA 60
QY 76 VSIYIINLVAAADFLPSGHIICSPRLINIRHPISKILSPVMTFPYFIGLSMISAISTER 135
DB 61 VSIYIINLVAAADFLPSGHIICSPRLINIRHPISKILSPVMTFPYFIGLSMISAISTER 120
QY 136 CLSIIMPPIWHCRPRYLSSVMCVLMAISLSTLEMMPCDPLFSGADSVWCETSDFT 195
DB 121 CLSIIMPPIWHCRPRYLSSVMCVLMAISLSTLEMMPCDPLFSGADSVWCETSDFT 180
QY 196 IAMLVFLCVLLCGSSVLVLRILCGSRKMPLTRLYTITLVVFLLCGLPFGIQWALFS 255
DB 181 IAMLVFLCVLLCGSSVLVLRILCGSRKMPLTRLYTITLVVFLLCGLPFGIQWALFS 240
QY 256 RIHLDMKVLFCFVHLVSIPLSALNSSANPIIYFVGSFRORONRKLVLRALODTPE 315
DB 241 RIHLDMKVLFCFVHLVSIPLSALNSSANPIIYFVGSFRORONRKLVLRALODTPE 300
QY 316 VDEGGGMLPOETLESGRLQ 337
DB 301 VDEGGGMLPOETLESGRLQ 322
RESULT 9

ADO44602
ID ADO44602 standard; protein; 322 AA.
XX
XX ADO44602;
AC
XX 29-JUL-2004 (first entry)
DE
XX Human HIT7213 protein.
XX
XX HIT7213; transgenic; G protein-coupled receptor; GPCR; ophthalmological;
KM cytostatic; nephrotoxic; antiinflammatory; dermatological; analgesic;
XX vulnerable; neuroprotective; human; receptor.
XX
XX Homo sapiens.
XX
XX WO2004039972-A1.
XX 13-MAY-2004.
XX 28-OCT-2003; 2003WO-JP013781.
XX 29-OCT-2002; 2002JP-00314141.
XX (TAKA) TAKEDA CHEM IND LTD.
XX
XX Kaisho Y, Watanabe T, Yasuhara Y, Mori I, Takeomi S;
XX WPI; 2004-376191/35.
XX N-PSDB; ADO44603.
XX
XX HIT7213 protein, encoded DNA and transgenic animals for clarifying
PT pathological mechanism, developing therapeutic methods and screening
PT preventives or remedies for related diseases e.g. cataract, cancer, and
PT dermatitis.
XX
XX Claim 3; SEQ ID NO 1; 161pp; Japanese.
XX
XX The invention relates to a non-human mammal that carries a DNA integrated
CC with a foreign HIT7213 or its mutant gene, or a part of it. The non-human
CC animal is particularly a rat. Such gene shows phenotypes of e.g. cataract
CC onset, transient skin rash and proliferation-promoting activity. The
CC foreign HIT7213 gene is a gene that encodes a G protein-coupled receptor
CC (GPCR) protein HIT7213. The protein, its encoded DNA and constructed
CC transgenic animals are useful for clarifying pathological mechanism,
CC developing therapeutic methods and screening preventives or remedies for
CC related diseases e.g. cataract, cancer, and dermatitis. The present
CC sequence represents a human HIT7213 protein.
XX
XX Sequence 322 AA;
SQ
Query Match 95.5%; Score 322; DB 8; Length 322;
Best Local Similarity 100.0%; Pred. No. 2.9e-309; Indels 0; Gaps 0;
Matches 322; Conservative 0; Mismatches 0;
QY 16 MDSTIPVLGTELPINGREBTPCYKOTLSFTGLTCTVSLVALTGNVAVMLLGCRRRNA 75
DB 1 MDSTIPVLGTELPINGREBTPCYKOTLSFTGLTCTVSLVALTGNVAVMLLGCRRRNA 60
QY 76 VSIYIINLVAAADFLPSGHIICSPRLINIRHPISKILSPVMTFPYFIGLSMISAISTER 135
DB 61 VSIYIINLVAAADFLPSGHIICSPRLINIRHPISKILSPVMTFPYFIGLSMISAISTER 120
QY 136 CLSIIMPPIWHCRPRYLSSVMCVLMAISLSTLEMMPCDPLFSGADSVWCETSDFT 195
DB 121 CLSIIMPPIWHCRPRYLSSVMCVLMAISLSTLEMMPCDPLFSGADSVWCETSDFT 180
QY 196 IAMLVFLCVLLCGSSVLVLRILCGSRKMPLTRLYTITLVVFLLCGLPFGIQWALFS 255
DB 181 IAMLVFLCVLLCGSSVLVLRILCGSRKMPLTRLYTITLVVFLLCGLPFGIQWALFS 240
QY 256 RIHLDMKVLFCFVHLVSIPLSALNSSANPIIYFVGSFRORONRKLVLRALODTPE 315
DB 241 RIHLDMKVLFCFVHLVSIPLSALNSSANPIIYFVGSFRORONRKLVLRALODTPE 300

QY 316 VDEGGWLPQETLEISGRLEQ 337
 DB 301 VDEGGWLPQETLEISGRLEQ 322

RESULT 10
 ID ADO29705 standard; protein; 322 AA.
 XX ADO29705;
 XX 29-JUL-2004 (first entry)
 DE Human GPCR MRGX3, SEQ ID NO:807.
 XX
 XX G protein-coupled receptor; GPCR; drug screening; diagnosis;
 KW transgenic mouse; neurological disorder; adrenal gland disorder;
 KW colon disorder; intestinal disorder; cardiovascular disorder;
 KW muscular disorder; blood disorder; immune disorder; bone disorder;
 KW joint disorder; metabolic disorder; nutritive disorder; cancer;
 KW kidney disorder; liver disorder; lung disorder; breast disorder;
 KW ovary disorder; uterus disorder; prostate disorder; testis disorder;
 KW skin disorder; stomach disorder; pancreas disorder; spleen disorder;
 KW thymus disorder; thyroid disorder; antiparkinsonian; antitumor;
 KW cytoskeletal; antiinflammatory; vasotropic; antianginal; antiarrhythmic;
 KW CNS; central nervous system; respiratory; antidiabetic; antidiabetic;
 KW virocid; hepatotropic; antibacterial; antianemic; antiseborrheic;
 KW dermatological; antitumor; antihypertensive; antidiabetic; antineoplastic;
 KW immunosuppressive; nephrotoxic; gene therapy; GPCR modulator; human;
 KW receptor.
 XX
 OS Homo sapiens.
 XX
 PN W02004040000-A2.
 XX 13-MAY-2004.
 XX
 PF 09-SEP-2003; 2003WO-US028226.
 XX
 PR 09-SEP-2002; 2002US-0409303P.
 PR 09-APR-2003; 2003US-0461329P.
 XX
 PA (PRIM-) PRIMAL INC.
 XX
 PI Galtanaris GA, Bergmann JF, Gragerov A, Hohmann J, Li P;
 PI Medisen L, McIlwain KL, Pavlova MN, Vassiliadis D, Zeng H;
 DR WPI, 2004-390329/36.
 XX N-PSDB; ADO30080.
 PT Novel mammalian G protein coupled receptors, useful for identifying
 PT compounds that modulates diagnosing and treating disease condition
 PT associated with GPCR dysfunction e.g. autoimmune diseases, angina
 PT pectoris, Parkinson's disease.
 XX
 XX Claim 151, SEQ ID NO 807; 542pp; English.

CC depression, diabetic neuropathy, Parkinson's disease or schizophrenia);
 CC disorders of the adrenal gland; disorders of the colon or intestine
 CC (e.g., Crohn's disease, diarrhoea, food poisoning or irritable bowel
 CC syndrome); cardiovascular disorders (e.g., angina, cardiac arrhythmia or
 CC myocardial infarction); muscular disorders; blood disorders (e.g.,
 CC anaemia or leukaemia); immune disorders (e.g., autoimmune disorders or
 CC AIDS); bone and joint disorders (e.g., osteoarthritis, rheumatoid
 CC arthritis, gout or osteoporosis); metabolic or nutritive disorders (e.g.,
 CC obesity, enzyme deficiency-related diseases or vitamin deficiency-related
 CC diseases); and disorders of the kidney, liver, lung, breast, ovary,
 CC uterus, prostate, testis, skin, stomach, pancreas, spleen, thymus and
 CC thyroid (e.g., cancers). The present sequence represents a GPCR of the
 CC invention. Note: The full sequence data for this patent did not form part
 CC of the printed specification; those sequences not shown were obtained in
 CC electronic format directly from WIPO at
 CC ftp.wipo.int/pub/published_pcr_sequences.

XX
 SQ Sequence 322 AA;
 Query Match 95.5%; Score 322; DB 8; Length 322;
 Best Local Similarity 100.0%; Pred. No. 2.9e-309;
 Matches 322; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 16 MDSTIPVLTGTLTPNGEETPCYKQTLSPFGLCTIVSLVATGNVAVLWLGCRMRNA 75
 DB 1 MDSTIPVLTGTLTPNGEETPCYKQTLSPFGLCTIVSLVATGNVAVLWLGCRMRNA 60
 QY 76 VSIYILNVADPLFLSGHIIICSPRLINIRHPIKSLISPMTPFPYIGLSMLAISTER 135
 DB 61 VSIYILNVADPLFLSHIICSPRLINIRHPIKSLISPMTPFPYIGLSMLAISTER 120
 QY 136 CISTIPMWHCRPRRYISSVVCVLLWALSILRSILEMFCDFLFGSADSVWCETSDPT 195
 DB 121 CISTIPMWHCRPRRYISSVVCVLLWALSILRSILEMFCDFLFGSADSVWCETSDPT 180
 QY 196 IAMLVFLCVLGGSSVLVAVIILGSRKMPTRLYVYVILTVYVFLCGPFGIQAALFS 255
 DB 181 IAMLVFLCVLGGSSVLVAVIILGSRKMPTRLYVYVILTVYVFLCGPFGIQAALFS 240
 QY 256 RIHLDKVLFCVHVLVSIIFLSALNSSANPIIYFFVGSFRORONKMLKVLQALDPTPE 315
 DB 241 RIHLDKVLFCVHVLVSIIFLSALNSSANPIIYFFVGSFRORONKMLKVLQALDPTPE 300

QY 316 VDEGGWLPQETLEISGRLEQ 337
 DB 301 VDEGGWLPQETLEISGRLEQ 322

RESULT 11
 ID ADF70481 standard; protein; 560 AA.
 XX ADF70481;
 AC ADF70481;
 XX
 DT 12-FEB-2004 (first entry)
 XX
 DB Orphan receptor ligand-related human protein SegID104.
 XX
 KW ligand; orphan receptor protein; fusion protein; fluorescent protein;
 KW cell expression; green fluorescent protein; GFP; GFP-1; wild-type GFP;
 KW GFPuv; Enhanced GFP; EGFP; human.
 XX
 OS Homo sapiens.
 XX
 PN W02003071272-A1.
 XX
 PD 28-AUG-2003.
 XX
 PF 21-FEB-2003; 2003WO-JP001901.
 XX
 PR 22-FEB-2002; 2002JP-00045728.
 PR 23-JUL-2002; 2002JP-00213949.
 PR 11-OCT-2002; 2002JP-00298237.

XX (TAKEDA CHEM IND LTD.
XX
XX Hinuma S, Fujii R, Ogi K, Komatsu H, Kawamata Y, Hosoya M;
XX WPI; 2003-697654/66.
XX DR N-PSDB; ADF70583.
XX
XX Transformation of cells with a fusion protein of an orphan receptor
XX PT protein with a fluorescent protein useful for identification of ligands
XX PT to the orphan receptor.
XX
XX Disclosure; SEQ ID NO 104; 594pp; Japanese.
XX
XX This invention relates to a novel method of identifying ligands to an
XX orphan receptor protein which comprises transforming cells with DNA
XX encoding a fusion protein of the orphan receptor with a fluorescent
XX protein, so that the fusion protein is expressed in the cells (or cell
XX membranes isolated from them) and contacting the cells with the potential
XX ligand to be tested. A suitable fluorescent protein for incorporation in
XX the fusion protein is green fluorescent protein (GFP), for example GFP-1,
XX wild-type GFP, GFPuv or Enhanced GFP (EGFP). The method is useful for the
XX identification of ligands binding to an orphan receptor protein.
XX
XX Sequence 560 AA:
SQ
Query Match 95.5%; Score 322; DB 7; Length 560;
Best Local Similarity 100.0%; Pred. No. 4.9e-309;
Matches 322; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 16 MDSTIPVLGTELPINGREETPCYKOTLSTFTGLTCIYSLVATGNNAVWMLGCRMRNA 75
DB 1 MDSTIPVLGTELPINGREETPCYKOTLSTFTGLTCIYSLVATGNNAVWMLGCRMRNA 60
QY 76 VSIYIILNVAADFLPSGHIICSPRLINIRHPISKLSVMTPTPYFTGLSMTSAISTER 135
DB 61 VSIYIILNVAADFLPSGHIICSPRLINIRHPISKLSVMTPTPYFTGLSMTSAISTER 120
QY 136 CSTIIMPWYHCRPRYLSSVWCYLWALSLRSTLEMMPCDPLFSGADSWCETSDFT 195
DB 121 CSTIIMPWYHCRPRYLSSVWCYLWALSLRSTLEMMPCDPLFSGADSWCETSDFT 180
QY 196 IAWLVPLCVLCCSSLVTLVRIILGSRKMPLTRLYVTLLTVLVFLGCPFGIQWALFS 255
DB 181 IAWLVPLCVLCCSSLVTLVRIILGSRKMPLTRLYVTLLTVLVFLGCPFGIQWALFS 240
QY 256 RIHLDMKVLFCFVHLVSIPLSALNSSANPIIYFFVGSFRQRONRKIKVLQRALQDTPR 315
DB 241 RIHLDMKVLFCFVHLVSIPLSALNSSANPIIYFFVGSFRQRONRKIKVLQRALQDTPR 300
QY 316 VDEGGGMLPQETLELSSRLQ 337
DB 301 VDEGGGMLPQETLELSSRLQ 322
RESULT 12
ADP29105
ID ADF29105 standard; protein; 322 AA.
AC ADF29105;
XX
XX 12-FEB-2004 (first entry)
XX DT
XX Human GPCR R-14.
XX DE
XX Human; R-14; GPCR; G protein coupled receptor;
XX trabeular meshwork tissue; ocular outflow; antagonist;
XX ocular hypotensive; drug screening; elevated intraocular pressure;
XX glaucoma; ophthalmological; receptor.
XX KM
XX Homo sapiens.
XX OS
XX
XX WO2003080659-A1.

XX 02-OCT-2003.
XX PD
XX 27-MAR-2003; 2003WO-CA000444.
XX PF
XX 27-MAR-2002; 2002US-0367513P.
XX FR
XX (THER-) THERATECHNOLOGIES INC.
XX PA
XX Peri KG, Mofelt S, Abran D;
XX PI
XX WPI; 2004-053019/05.
XX DR N-PSDB; ADF29104.
XX
XX New substantially pure R-14 polypeptide, useful as drug targets for
XX PT lowering intraocular pressure and for treating condition such as
XX PT glaucoma.
XX
XX Claim 21; SEQ ID NO 2; 85pp; English.
XX PS
XX The invention relates to a human GPCR (G protein coupled receptor)
XX CC designated R-14 (ADP29105), nucleic acids encoding it (ADP29104), and R-
XX CC 14 peptide antagonists (ADP29106-ADP29108). The invention also
XX CC encompasses vectors and host cells comprising R-14 nucleic acids, and
XX CC methods of screening for R-14 antagonists. The R-14 receptor is expressed
XX CC in trabecular meshwork tissue, and is associated with a role in ocular
XX CC outflow. Inhibition of the receptor results in a reduction of basal
XX CC intraocular pressure, making the R-14 receptor a useful target for
XX CC screening for ocular hypotensive drugs. The R-14 peptide antagonists are
XX CC useful for reducing intraocular pressure for the treatment of conditions
XX CC associated with elevated intraocular pressure such as glaucoma and
XX CC related conditions. The R-14 receptor, and host cells expressing an R-14
XX CC polynucleotide, may be used in screening for R-14 receptor antagonists.
XX The present sequence is related to the invention.
XX
XX Sequence 322 AA:
SQ
Query Match 82.2%; Score 277; DB 8; Length 322;
Best Local Similarity 100.0%; Pred. No. 8.6e-265;
Matches 277; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 61 AAVLWMLGCRMRNAVSIYILNVAADFLPSGHIICSPRLINIRHPISKLSVMTPTFP 120
DB 46 AAVLWMLGCRMRNAVSIYILNVAADFLPSGHIICSPRLINIRHPISKLSVMTPTFP 105
QY 121 YFTGLSMTLSISTERCLSTIMPWYHCRPRYLSSVWCYLWALSLRSTLEMMPCDPLF 180
DB 106 YFTGLSMTLSISTERCLSTIMPWYHCRPRYLSSVWCYLWALSLRSTLEMMPCDPLF 165
QY 181 SGADSWCETSDFTTAMLVFLCVLCCSSLVTLVRIILGSRKMPLTRLYVTLLTVLVF 240
DB 166 SGADSWCETSDFTTAMLVFLCVLCCSSLVTLVRIILGSRKMPLTRLYVTLLTVLVF 225
QY 241 LILGCPFGIQWALFSRIHLDMKVLFCFVHLVSIPLSALNSSANPIIYFFVGSFRQRONRQ 300
DB 226 LILGCPFGIQWALFSRIHLDMKVLFCFVHLVSIPLSALNSSANPIIYFFVGSFRQRONRQ 285
QY 301 NLKLVTLQRALQDTPRVDGEGGMLPQETLELSSRLQ 337
DB 286 NLKLVTLQRALQDTPRVDGEGGMLPQETLELSSRLQ 322
RESULT 13
AAB14846
ID AAB14846 standard; protein; 322 AA.
AC AAB14846;
XX
XX 19-DEC-2000 (first entry)
XX DT
XX Human novel G-protein coupled receptor #1.
XX DE
XX Human; novel G-protein coupled receptor; signal transduction;
XX KM

KW disease diagnosis; drug screening; disease therapy.
XX Homo sapiens.
XX WO200040724-A1.
XX
XX 13-JUL-2000.
XX
XX 04-JAN-2000; 2000WO-US000052.
XX
XX 04-JAN-1999; 99US-0114666P.
PR 14-JAN-1999; 99US-0115828P.
XX
XX (LEXI-) LEXICON GENETICS INC.
XX
XX Nehls M, Wattler F;
XX WPI; 2000-465986/40.
DR N-PSDB; AAA70342.
XX
XX New polynucleotides encoding novel G-protein coupled receptors useful for
PT diagnosis, drug screening, clinical trial monitoring and for the
PT treatment of physiological or behavioral disorders.
XX
XX Claim 1; Page 53-54; 61pp; English.
XX
XX The present sequence is the protein sequence for a novel human G-protein
CC coupled receptor (NGPCR). These proteins are involved in signal
CC transduction pathways in many cases. The protein contains seven
CC transmembrane domains, and is expressed in human testis, mammary gland
CC and salivary gland tissue. The protein, its gene, agonists, antagonists
CC and antibodies can be used to diagnose and treat diseases associated with
CC the inappropriate expression or expression of mutant versions of the
CC protein, for screening for drugs which can be used in the same manner,
CC and for elucidating the function of the protein
XX
XX
SQ Sequence 322 AA;
Query Match 66.5%; Score 224; DB 3; Length 322;
Best Local Similarity 100.0%; Pred. No. 2.1e-212;
Matches 224; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 98 SPLRLINIRHPISKLSVMTFPYPIGSLMSAISTERCLSLTWPIMWYHCRPRYLSSVM 157
DB 83 SPLRLINIRHPISKLSVMTFPYPIGSLMSAISTERCLSLTWPIMWYHCRPRYLSSVM 142
QY 158 CVLMLALSILSLSEMMFCDFLFGSADSVWCETSDPTITAMLVFLCVLTCSSLVLVARI 217
DB 143 CVLMLALSILSLSEMMFCDFLFGSADSVWCETSDPTITAMLVFLCVLTCSSLVLVARI 202
QY 218 LCGSRKMPRLRYVITLTVLVEFLCGPFGIQMALFSRIHDMKVLFGCHLVLSIFLSA 277
DB 203 LCGSRKMPRLRYVITLTVLVEFLCGPFGIQMALFSRIHDMKVLFGCHLVLSIFLSA 262
QY 278 LNSSANPIIYFVGSFRORONRQNLKVLQRLADOTPEVDEGGG 321
DB 263 LNSSANPIIYFVGSFRORONRQNLKVLQRLADOTPEVDEGGG 306
RESULT 14
AA90762
ID AAY90762 standard; protein; 322 AA.
XX
XX AAY90762;
XX
XX 18-AUG-2000 (first entry)
XX
XX Human G protein-coupled receptor hHR77213V SEQ ID NO:2.
XX
XX Human; G protein-coupled receptor; hippocampus; diagnosis; screening;
KM genetic disease; cellular function regulation.
XX
OS Homo sapiens.

XX
XX WO200020455-A1.
XX
XX 13-APR-2000.
XX
XX 30-SEP-1999; 99WO-JP005366.
XX
XX 01-OCT-1998; 98JP-00279535.
XX
XX (TAKE) TAKEDA CHEM IND LTD.
XX
XX Watanabe T, Terao Y, Matsui H;
XX WPI; 2000-303747/26.
DR N-PSDB; AAA29812.
XX
XX Human-derived G protein-coupled protein and encoding nucleic acid, useful
PT e.g. in determining ligands and treatment of diseases associated with
PT dysfunction of the protein.
XX
XX Claim 2; Page 92-93; 97pp; Japanese.
XX
XX The present sequence represents a human-derived G protein-coupled protein
CC designated hHR77213V, which is isolated from the human hippocampus. The G
CC protein-coupled receptor can be used for preventing, treating and
CC diagnosing genetic diseases associated with G protein-coupled protein,
CC and for regulating cellular functions. The protein can be used to prevent
CC and treat disorders associated with G protein-coupled protein gene
CC dysfunction. It can also be used to identify G protein-coupled protein
CC ligands and generating antibodies and antisera against the protein. It is
CC also useful in constructing recombinant receptor protein expression
CC systems, developing receptor-binding assay systems and screening drug
CC candidates, and can be used as a probe in the genetic diagnosis of G
CC protein-coupled protein disorders
XX
XX
SQ Sequence 322 AA;
Query Match 65.6%; Score 221; DB 3; Length 322;
Best Local Similarity 99.7%; Pred. No. 1.9e-209;
Matches 321; Conservative 0; Mismatches 1; Indels 0; Gaps 0;
QY 16 MDSTIPVGTSLTPNGRETPCYKQTLSPFGTLCTIVSLATGNAAVLMGLCRMNRNA 75
DB 1 MDSTIPVGTSLTPNGRETPCYKQTLSPFGTLCTIVSLATGNAAVLMGLCRMNRNA 60
QY 76 VSIYILNVAADFLFLSGHITCSPLRLINIRHPISKLSVMTFPYPIGSLMSAISTER 135
DB 61 VSIYILNVAADFLFLSGHITCSPLRLINIRHPISKLSVMTFPYPIGSLMSAISTER 120
QY 136 CLSTIMPIWYHCRPRYLSSVMCVLMLALSILSLSEMMFCDFLFGSADSVWCETSDPT 195
DB 121 CLSTIMPIWYHCRPRYLSSVMCVLMLALSILSLSEMMFCDFLFGSADSVWCETSDPT 180
QY 196 IAMLVFLCVLTCSSLVLVARIILCGSRKMPRLRYVITLTVLVEFLCGPFGIQMALFS 255
DB 181 IAMLVFLCVLTCSSLVLVARIILCGSRKMPRLRYVITLTVLVEFLCGPFGIQMALFS 240
QY 256 RIHDMKVLFGCHLVLSIFLSALNSSANPIIYFVGSFRORONRQNLKVLQRLADOTPE 315
DB 241 RIHDMKVLFGCHLVLSIFLSALNSSANPIIYFVGSFRORONRQNLKVLQRLADOTPE 300
QY 316 VDEGGWLPQETLISGRLAQ 337
DB 301 VDEGGWLPQETLISGRLAQ 322
RESULT 15
AD116994
ID AD116994 standard; protein; 322 AA.
XX
XX AD116994;
XX
XX 15-APR-2004 (first entry)
XX
XX

XX	DE	Human NOVX protein homologue SegID 530.
XX	KW	human; NOVX; cardiomyopathy; atherosclerosis; cancer; diabetes;
XX	KW	inflammation; autoimmune disorder; allergy; blood disorder;
XX	KW	acquired immunodeficiency syndrome; AIDS; obesity; asthma;
XX	KW	immunoglobulin (Ig) A nephropathy; cirrhosis; arthritis;
XX	KW	Alzheimer's disease; infection; str.
XX	OS	Homo sapiens.
XX	PN	WO200268649-A2.
XX	PD	06-SEP-2002.
XX	PP	31-JAN-2002; 2002WO-US002785.
XX	PR	31-JAN-2001; 2001US-0265395P.
XX	PR	31-JAN-2001; 2001US-0265412P.
XX	PR	31-JAN-2001; 2001US-0265514P.
XX	PR	31-JAN-2001; 2001US-0265517P.
XX	PR	02-FEB-2001; 2001US-0266406P.
XX	PR	05-FEB-2001; 2001US-0266767P.
XX	PR	07-FEB-2001; 2001US-0266975P.
XX	PR	07-FEB-2001; 2001US-0267057P.
XX	PR	08-FEB-2001; 2001US-0267459P.
XX	PR	09-FEB-2001; 2001US-0267823P.
XX	PR	15-FEB-2001; 2001US-0268974P.
XX	PR	26-FEB-2001; 2001US-0271664P.
XX	PR	27-FEB-2001; 2001US-0271839P.
XX	PR	27-FEB-2001; 2001US-0271855P.
XX	PR	02-MAR-2001; 2001US-0272788P.
XX	PR	02-MAR-2001; 2001US-0273046P.
XX	PR	14-MAR-2001; 2001US-0275925P.
XX	PR	14-MAR-2001; 2001US-0275947P.
XX	PR	14-MAR-2001; 2001US-0275950P.
XX	PR	14-MAR-2001; 2001US-0275989P.
XX	PR	15-MAR-2001; 2001US-0276448P.
XX	PR	15-MAR-2001; 2001US-0276450P.
XX	PR	16-MAR-2001; 2001US-0276397P.
XX	PR	16-MAR-2001; 2001US-0276768P.
XX	PR	20-MAR-2001; 2001US-0278652P.
XX	PR	26-MAR-2001; 2001US-0278775P.
XX	PR	26-MAR-2001; 2001US-0278778P.
XX	PR	29-MAR-2001; 2001US-0279882P.
XX	PR	29-MAR-2001; 2001US-0279884P.
XX	PR	30-MAR-2001; 2001US-0280147P.
XX	PR	11-APR-2001; 2001US-0282992P.
XX	PR	11-APR-2001; 2001US-0283083P.
XX	PR	23-APR-2001; 2001US-0285133P.
XX	PR	23-APR-2001; 2001US-0285749P.
XX	PR	03-MAY-2001; 2001US-0288327P.
XX	PR	03-MAY-2001; 2001US-0288504P.
XX	PR	29-MAY-2001; 2001US-0294047P.
XX	PR	30-MAY-2001; 2001US-0294473P.
XX	PR	08-JUN-2001; 2001US-0296964P.
XX	PR	18-JUN-2001; 2001US-0298959P.
XX	PR	19-JUN-2001; 2001US-0299324P.
XX	PR	13-AUG-2001; 2001US-0312020P.
XX	PR	16-AUG-2001; 2001US-0312889P.
XX	PR	16-AUG-2001; 2001US-0312908P.
XX	PR	21-AUG-2001; 2001US-0313390P.
XX	PR	28-AUG-2001; 2001US-0315470P.
XX	PR	31-AUG-2001; 2001US-0316447P.
XX	PR	07-SEP-2001; 2001US-0318115P.
XX	PR	07-SEP-2001; 2001US-0318118P.
XX	PR	12-SEP-2001; 2001US-0318740P.
XX	PR	19-SEP-2001; 2001US-0323379P.
XX	PR	18-OCT-2001; 2001US-0330245P.
XX	PR	18-OCT-2001; 2001US-0330308P.
XX	PR	14-NOV-2001; 2001US-0332701P.
XX	PA	(CURA-) CURAGEN CORP.

XX Tchervnev VT, Sopytek KA, Zernusen BD, Patutrafjan M, Shinkets RA;
PI Li L, Ganggoli EA, Padigar M, Anderson DW, Rastelli L, Miller CB;
PI Getlach VU, Taupier RJ, Gusev VY, Colman SD, Wolencik A, Pena CA;
PI Furtak K, Grose WM, Albrook JP, Lepley DM, Rieger DK, Burgess CE;
XX WPI; 2002-706998/76.
DR
XX
XX New NOVX polypeptides and nucleic acids, useful for preventing or
PT treating NOVX-associated disorders, e.g. cancer, cardiomyopathy,
PT atherosclerosis, or diabetes, and in chromosome mapping, tissue typing or
PT pharmacogenomics.

Disclosure; SEQ ID NO 530; 1498bp; English

This invention relates to a novel nucleic acids, and encoded polypeptides thereof, which have properties related to the stimulation of biochemical or physiological responses in a cell, tissue, organ or organism. Specifically, it refers to the use of biologically active fragments for diagnostic and prognostic assays and furthermore in the treatment of diverse pathological conditions. The present invention describes novel human and murine NOX proteins, as well as methods to modulate their expression using antisense oligos, ribozymes and peptide nucleic acids. The NOX polypeptides, polynucleotides and antibodies are useful in treating or preventing NOX-associated disorders, e.g. cardiomyopathy, atherosclerosis, cancer and diabetes. Furthermore, they may be used in treating or preventing diseases such as inflammation, autoimmune disorders, allergies, blood disorders, acquired immunodeficiency syndrome (AIDS), obesity, asthma, immunoglobulin (Ig) nephropathy, cirrhosis, arthritis, Alzheimer's disease, infections, stroke, muscular dystrophy and epilepsy. Accordingly, these molecules have many activities including cytoskeletal, cardiac, anti-inflammatory, immunosuppressive, antiallergic, haemostatic, anti-HIV, antidiabetic, antiarteriosclerotic, anorectic, antiaesthetic, nephrotoxic, antiarthritic, hepatotropic, neuroprotective, nootropic, antibacterial, virucide, antiparasitic, relaxant and anticonvulsant. In addition, they are useful in screening assays to identify small molecules that modulate or inhibit, for example, neurogenesis, wound healing and angiogenesis. The nucleic acids are also used as in chromosome mapping, tissue typing, preventive medicine and pharmacogenomics. This polypeptide is a homologue of a human NOX protein of the invention.

SQ Sequence 322 AA;

Query Match	Score	DB	Length
65.6%	221	5	322

Matches 321; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

16 MDSTIPVLGTELTPIINGREETPCYKQTLSTFTGLTCIVSLVALTGNVVLWLLGCRMRNA 75

Db 1 MDSTIPVLGTELTPIINGREETPCYKQTLSTFTGLTCIVSLVALTGNVAVLWLLGCRM RNA 60

Qy 76 VSIYIINLVADFLFLSGHIICSPRLINIRHPISKILSPVMTFFPYFIGLSMLSAISTER 135

Db 61 VSIYIINLVAADFLFLSGHIICSPRLINIRHPISKILSPVMTFFPYFIGLSMLSAISTER 120

Qy 136 CLSILWPIWYHCRPRYLSSVMCVLLWALSILRSILEWMFCDPLFGADSVWCETSDFIT 195

Db 121 CLSILWPIWYHCRPRYLSVCMVLLWALSLRSILEWMFCDFLFGANSVWCETSDFIT 180

QY 196 IAWLVFLCVLGGSSLVLLVRILCGSRKMPLTRLYVTILLTVLVFLCGLPFGIQWALFS 255

Db 181 IAWLVFLCVVLCGSSLVLLVRILCGSRKMPTRLVTILTLVLFLLCGLPFGIQWALFS 240

256 RIHLDWKVLFCVHVLVSI FLSALNSSANPI IYFFVGSFRQRQRNLKVLQRALQDTPE 3155

Db 241 RIHLDWKVLECHVHLSIFLSALNSSANPIIYFFVGSRQRQRNLKVLQRALQDTPE 300

QY 316 VDEGGWLPQETLELSGSRLQ 337

Db 301 VDEGGGWLPEETLELSGSRLAQ 322

(CURA-) CURAGEN CORP.

Mon Nov 1 12:58:47 2004

us-09-867-570-2.olg.rag

Page 12

Search completed: October 27, 2004, 08:48:14
Job time : 156 secs

```

; APPLICANT: Shi-Hsiang, Shen
; TITLE OF INVENTION: G Protein-Coupled Receptors from the Rat and Human
; FILE REFERENCE: 81823/268117
; CURRENT APPLICATION NUMBER: US/09/254,227A
; CURRENT FILING DATE: 1999-03-03
; NUMBER OF SEQ ID NOS: 22
; SOFTWARE: Patentin version 3.0
; SEQ ID NO 5
; LENGTH: 322
; TYPE: PRT
; ORGANISM: Homo sapiens
US-09-254-227A-5

```

```

Query Match          25.5%; Score 86; DB 4; Length 322;
Best Local Similarity 100.0%; Pred. No. 6.5e-74;
Matches 86; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

```

```

Qy      204 VLGGSSIVLVRIICGSRKMPRLTYTITLTVVFLICGIPFGIOWALFGRHLDWKV 263
Db      189 VLGGSSIVLVRIICGSRKMPRLTYTITLTVVFLICGIPFGIOWALFGRHLDWKV 248
Qy      264 LFCVHLVSIPLSALNSSANPIIYF 289
Db      249 LFCVHLVSIPLSALNSSANPIIYF 274

```

```

RESULT 3
US-09-254-227A-9
; Sequence 9, Application US/09254227A
; Patent No. 6696257
; GENERAL INFORMATION:
; APPLICANT: Ahmad, Sultan
; APPLICANT: Banville, Denis
; APPLICANT: Fortin, Yves
; APPLICANT: Lembo, Paola
; APPLICANT: O'Donnell, Dajan
; APPLICANT: Shi-Hsiang, Shen
; TITLE OF INVENTION: G Protein-Coupled Receptors from the Rat and Human
; FILE REFERENCE: 81823/268117
; CURRENT APPLICATION NUMBER: US/09/254,227A
; CURRENT FILING DATE: 1999-03-03
; NUMBER OF SEQ ID NOS: 22
; SOFTWARE: Patentin version 3.0
; SEQ ID NO 9
; LENGTH: 322
; TYPE: PRT
; ORGANISM: Homo sapiens
US-09-254-227A-9

```

```

Query Match          14.8%; Score 50; DB 4; Length 322;
Best Local Similarity 100.0%; Pred. No. 1.6e-39;
Matches 50; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

```

```

Qy      263 VLFCVHLVSIPLSALNSSANPIIYFVGGFRORONRQNLKVIQRALOD 312
Db      248 VLFCVHLVSIPLSALNSSANPIIYFVGGFRORONRQNLKVIQRALOD 297

```

RESULT 4

```

US-09-254-227A-7
; Sequence 7, Application US/09254227A
; Patent No. 6696257
; GENERAL INFORMATION:
; APPLICANT: Ahmad, Sultan
; APPLICANT: Banville, Denis
; APPLICANT: Fortin, Yves
; APPLICANT: Lembo, Paola
; APPLICANT: O'Donnell, Dajan
; APPLICANT: Shi-Hsiang, Shen
; TITLE OF INVENTION: G Protein-Coupled Receptors from the Rat and Human
; FILE REFERENCE: 81823/268117
; CURRENT APPLICATION NUMBER: US/09/254,227A
; CURRENT FILING DATE: 1999-03-03

```

```

; NUMBER OF SEQ ID NOS: 22
; SOFTWARE: Patentin version 3.0
; SEQ ID NO 7
; LENGTH: 322
; TYPE: PRT
; ORGANISM: Homo sapiens
US-09-254-227A-7

```

```

Query Match          13.1%; Score 44; DB 4; Length 322;
Best Local Similarity 100.0%; Pred. No. 8.4e-34;
Matches 44; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

```

```

Qy      294 RORONRQNLKVIQRALODTPEVDGGMVLPQETLHSGSRLEQ 337
Db      279 RORONRQNLKVIQRALODTPEVDGGMVLPQETLHSGSRLEQ 322

```

```

RESULT 5
US-09-254-227A-11
; Sequence 11, Application US/09254227A
; Patent No. 6696257
; GENERAL INFORMATION:
; APPLICANT: Ahmad, Sultan
; APPLICANT: Banville, Denis
; APPLICANT: Fortin, Yves
; APPLICANT: Lembo, Paola
; APPLICANT: O'Donnell, Dajan
; APPLICANT: Shi-Hsiang, Shen
; TITLE OF INVENTION: G Protein-Coupled Receptors from the Rat and Human
; FILE REFERENCE: 81823/268117
; CURRENT APPLICATION NUMBER: US/09/254,227A
; CURRENT FILING DATE: 1999-03-03
; NUMBER OF SEQ ID NOS: 22
; SOFTWARE: Patentin version 3.0
; SEQ ID NO 11
; LENGTH: 322
; TYPE: PRT
; ORGANISM: Homo sapiens
US-09-254-227A-11

```

```

Query Match          12.2%; Score 41; DB 4; Length 322;
Best Local Similarity 100.0%; Pred. No. 6.2e-31;
Matches 41; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

```

```

Qy      209 SSVLVVRIICGSRKMPRLTYTITLTVVFLICGIPFGI 249
Db      194 SSVLVVRIICGSRKMPRLTYTITLTVVFLICGIPFGI 234

```

```

RESULT 6
US-09-254-227A-13
; Sequence 13, Application US/09254227A
; Patent No. 6696257
; GENERAL INFORMATION:
; APPLICANT: Ahmad, Sultan
; APPLICANT: Banville, Denis
; APPLICANT: Fortin, Yves
; APPLICANT: Lembo, Paola
; APPLICANT: O'Donnell, Dajan
; APPLICANT: Shi-Hsiang, Shen
; TITLE OF INVENTION: G Protein-Coupled Receptors from the Rat and Human
; FILE REFERENCE: 81823/268117
; CURRENT APPLICATION NUMBER: US/09/254,227A
; CURRENT FILING DATE: 1999-03-03
; NUMBER OF SEQ ID NOS: 22
; SOFTWARE: Patentin version 3.0
; SEQ ID NO 13
; LENGTH: 322
; TYPE: PRT
; ORGANISM: Homo sapiens
US-09-254-227A-13

```

Query Match 12.2%; Score 41; DB 4; Length 322;

Best Local Similarity 100.0%; Pred. No. 6.2e-31;
Matches 41; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 209 SSILVLRICGSRKMPILTYTITLTVAVPLICGPFQI 249
DB 194 SSILVLRICGSRKMPILTYTITLTVAVPLICGPFQI 234

RESULT 7

US-09-254-227A-1
; Sequence 1, Application US/09254227A
; Patent No. 6696257
; GENERAL INFORMATION:
; APPLICANT: Ahmad, Sultan
; APPLICANT: Banville, Denis
; APPLICANT: Portin, Yves
; APPLICANT: Lembo, Paola
; APPLICANT: O'Donnell, Dajan
; APPLICANT: Shi-Hsiang, Shen
; TITLE OF INVENTION: G Protein-Coupled Receptors from the Rat and Human
; FILE REFERENCE: 81823/268117
; CURRENT APPLICATION NUMBER: US/09/254,227A
; CURRENT FILING DATE: 1999-03-03
; NUMBER OF SEQ ID NOS: 22
; SOFTWARE: Patentin version 3.0
; SEQ ID NO 1
; LENGTH: 337
; TYPE: PRT
; ORGANISM: rat
US-09-254-227A-1

Query Match 3.3%; Score 11; DB 4; Length 337;
Best Local Similarity 100.0%; Pred. No. 0.029;
Matches 11; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 128 LSAISTERCLS 138
DB 131 LSAISTERCLS 141

RESULT 8

US-09-489-039A-10936
; Sequence 10936, Application US/09489039A
; Patent No. 6610836
; GENERAL INFORMATION:
; APPLICANT: Gary Breton et. al
; TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO KLEBSIELLA
; FILE REFERENCE: 2709.2004001
; CURRENT APPLICATION NUMBER: US/09/489,039A
; CURRENT FILING DATE: 2000-01-27
; PRIOR APPLICATION NUMBER: US 60/117,747
; PRIOR FILING DATE: 1999-01-29
; NUMBER OF SEQ ID NOS: 14342
; SEQ ID NO 10936
; LENGTH: 773
; TYPE: PRT
; ORGANISM: Klebsiella pneumoniae
US-09-489-039A-10936

Query Match 3.0%; Score 10; DB 4; Length 773;
Best Local Similarity 100.0%; Pred. No. 0.56;
Matches 10; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 271 VSIFLSALNS 280
DB 523 VSIFLSALNS 532

RESULT 9
US-09-255-368-8
; Sequence 8, Application US/09255368
; Patent No. 6262246

; GENERAL INFORMATION:
; APPLICANT: Gerald, Christophe P.G.
; APPLICANT: Jones, Kenneth A.
; APPLICANT: Bonini, James A.

; APPLICANT: Borowsky, Beth
; TITLE OF INVENTION: DNA Encoding Mammalian Neuropeptide RF (NPFF) Receptors
; FILE REFERENCE: 1795/57155-A
; CURRENT APPLICATION NUMBER: US/09/255,368
; CURRENT FILING DATE: 1999-02-22
; EARLIER APPLICATION NUMBER: 09/161,113
; EARLIER FILING DATE: 1998-09-25
; NUMBER OF SEQ ID NOS: 42
; SOFTWARE: Patentin Ver. 2.0 - beta
; SEQ ID NO 8
; LENGTH: 430
; TYPE: PRT
; ORGANISM: Homo sapiens
US-09-255-368-8

Query Match 2.7%; Score 9; DB 3; Length 430;
Best Local Similarity 100.0%; Pred. No. 2.9;
Matches 9; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 279 NSSANPIIY 287
DB 321 NSSANPIIY 329

RESULT 10
US-09-405-558-8
; Sequence 8, Application US/09405558A
; Patent No. 6709831

; GENERAL INFORMATION:
; APPLICANT: Gerald, Christophe P.G.
; APPLICANT: Jones, Kenneth A.
; APPLICANT: Bonini, James A.
; APPLICANT: Borowsky, Beth E.
; APPLICANT: Craig, Douglas A.
; TITLE OF INVENTION: DNA Encoding Mammalian Neuropeptide RF (NPFF) Receptors
; FILE REFERENCE: 57155-C
; CURRENT APPLICATION NUMBER: US/09/405,558A
; CURRENT FILING DATE: 1999-09-24
; EARLIER APPLICATION NUMBER: 09/161,113
; EARLIER FILING DATE: 1998-09-25
; EARLIER APPLICATION NUMBER: 09/255,368
; EARLIER FILING DATE: 1999-02-22
; NUMBER OF SEQ ID NOS: 71
; SOFTWARE: Patentin Ver. 2.0 - beta
; SEQ ID NO 8
; LENGTH: 430
; TYPE: PRT
; ORGANISM: Homo sapiens
US-09-405-558-8

Query Match 2.7%; Score 9; DB 4; Length 430;
Best Local Similarity 100.0%; Pred. No. 2.9;
Matches 9; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 279 NSSANPIIY 287
DB 321 NSSANPIIY 329

RESULT 11
US-08-978-404B-31
; Sequence 31, Application US/08978404B
; Patent No. 5968782
; GENERAL INFORMATION:
; APPLICANT: Stevens, Richard L.
; TITLE OF INVENTION: MAST CELL PROTEASE THAT CLEAVES
; TITLE OF INVENTION: FIBRINOGEN

NUMBER OF SEQUENCES: 74
CORRESPONDENCE ADDRESS:
ADDRESSEE: Wolf, Greenfield & Sacks, P.C.
STREET: 600 Atlantic Avenue
CITY: Boston
STATE: MA
COUNTRY: U.S.A.
ZIP: 02210-2211
COMPUTER READABLE FORM:
MEDIUM TYPE: Diskette
COMPUTER: IBM Compatible
OPERATING SYSTEM: DOS
SOFTWARE: FastSeq for Windows Version 2.0
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/978,404B
FILING DATE: 25-NOV-97
CLASSIFICATION: 435
PRIOR APPLICATION DATA:
APPLICATION NUMBER: 60/032,354
FILING DATE: 04-DEC-1996
ATTORNEY/AGENT INFORMATION:
NAME: Plumer, Elizabeth R.
REGISTRATION NUMBER: 36,637
REFERENCE/DOCKET NUMBER: B0801/7090
TELECOMMUNICATION INFORMATION:
TELEPHONE: 617-720-3500
TELEFAX: 617-720-2441
TELEX:
INFORMATION FOR SEQ ID NO: 31:
SEQUENCE CHARACTERISTICS:
LENGTH: 21 amino acids
TYPE: amino acid
STRANDEDNESS: single
TOPOLOGY: linear
MOLECULE TYPE: No. 5968782e
US-08-978-404B-31

Query Match 2.4%; Score 8; DB 2;
Best Local Similarity 100.0%; Pred. No. 1.6;
Matches 8; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 160 LLMALSL 167
DB 8 LLMALSL 15

RESULT 12
US-09-016-366A-15
Sequence 15, Application US/09016366A
Patent No. 5955431
GENERAL INFORMATION:
APPLICANT: Stevens, Richard L.
TITLE OF INVENTION: MAST CELL PROTEASE PEPTIDE
TITLE OF INVENTION: INHIBITORS
NUMBER OF SEQUENCES: 65
CORRESPONDENCE ADDRESS:
ADDRESSEE: Wolf, Greenfield & Sacks, P.C.
STREET: 600 Atlantic Avenue
CITY: Boston
STATE: MA
COUNTRY: U.S.A.
ZIP: 02210-2211
COMPUTER READABLE FORM:
MEDIUM TYPE: Diskette
COMPUTER: IBM Compatible
OPERATING SYSTEM: DOS
SOFTWARE: FastSeq for Windows Version 2.0
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/09/016,366A
FILING DATE: January 30, 1998
CLASSIFICATION: 530
PRIOR APPLICATION DATA:

APPLICATION NUMBER: 60/037,090
FILING DATE: 05-FEB-1997
ATTORNEY/AGENT INFORMATION:
NAME: Plumer, Elizabeth R.
REGISTRATION NUMBER: 36,637
REFERENCE/DOCKET NUMBER: B0801/7093
TELECOMMUNICATION INFORMATION:
TELEPHONE: 617-720-3500
TELEFAX: 617-720-2441
TELEX:
INFORMATION FOR SEQ ID NO: 15:
SEQUENCE CHARACTERISTICS:
LENGTH: 276 amino acids
TYPE: amino acid
STRANDEDNESS: single
TOPOLOGY: linear
MOLECULE TYPE: protein
US-09-016-366A-15

Query Match 2.4%; Score 8; DB 2;
Best Local Similarity 100.0%; Pred. No. 18;
Matches 8; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 160 LLMALSL 167
DB 8 LLMALSL 15

RESULT 13
US-08-978-404B-21
Sequence 21, Application US/08978404B
Patent No. 5968782
GENERAL INFORMATION:
APPLICANT: Stevens, Richard L.
TITLE OF INVENTION: MAST CELL PROTEASE THAT CLEAVES
TITLE OF INVENTION: FIBRINOGEN
NUMBER OF SEQUENCES: 74
CORRESPONDENCE ADDRESS:
ADDRESSEE: Wolf, Greenfield & Sacks, P.C.
STREET: 600 Atlantic Avenue
CITY: Boston
STATE: MA
COUNTRY: U.S.A.
ZIP: 02210-2211
COMPUTER READABLE FORM:
MEDIUM TYPE: Diskette
COMPUTER: IBM Compatible
OPERATING SYSTEM: DOS
SOFTWARE: FastSeq for Windows Version 2.0
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/978,404B
FILING DATE: 25-NOV-97
CLASSIFICATION: 435
PRIOR APPLICATION DATA:
APPLICATION NUMBER: 60/032,354
FILING DATE: 04-DEC-1996
ATTORNEY/AGENT INFORMATION:
NAME: Plumer, Elizabeth R.
REGISTRATION NUMBER: 36,637
REFERENCE/DOCKET NUMBER: B0801/7090
TELECOMMUNICATION INFORMATION:
TELEPHONE: 617-720-3500
TELEFAX: 617-720-2441
TELEX:
INFORMATION FOR SEQ ID NO: 21:
SEQUENCE CHARACTERISTICS:
LENGTH: 276 amino acids
TYPE: amino acid
STRANDEDNESS: single
TOPOLOGY: linear
MOLECULE TYPE: No. 5968782e
US-08-978-404B-21

Query Match 2.4%; Score 8; DB 2; Length 276;
 Best Local Similarity 100.0%; Pred. No. 18;
 Matches 8; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 160 LLMALSL 167
 |||||
 Db 8 LLMALSL 15

RESULT 14
 US-09-134-001C-3801
 ; Sequence 3801, Application US/09134001C
 ; Patent No. 6380370
 ; GENERAL INFORMATION:
 ; APPLICANT: Lynn Doucette-Stamm et al
 ; TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO STAPHYLOCOCCUS
 ; FILE REFERENCE: GTC-007
 ; CURRENT APPLICATION NUMBER: US/09/134,001C
 ; CURRENT FILING DATE: 1998-08-13
 ; PRIOR APPLICATION NUMBER: US 60/064,964
 ; PRIOR FILING DATE: 1997-11-08
 ; PRIOR APPLICATION NUMBER: US 60/055,779
 ; PRIOR FILING DATE: 1997-08-14
 ; NUMBER OF SEQ ID NOS: 5674
 ; SEQ ID NO 3801
 ; LENGTH: 428
 ; TYPE: PRT
 ; ORGANISM: Staphylococcus epidermidis
 US-09-134-001C-3801

Query Match 2.4%; Score 8; DB 3; Length 428;
 Best Local Similarity 100.0%; Pred. No. 26;
 Matches 8; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 322 WLPQETLE 329
 |||||
 Db 79 WLPQETLE 86

RESULT 15
 US-09-255-368-2
 ; Sequence 2, Application US/09255368
 ; Patent No. 6262246
 ; GENERAL INFORMATION:
 ; APPLICANT: Gerald, Christophe P.G.
 ; APPLICANT: Jones, Kenneth A.
 ; APPLICANT: Bonini, James A.
 ; APPLICANT: Borowsky, Beth
 ; TITLE OF INVENTION: DNA Encoding Mammalian Neuropeptide FF (NPFF) Receptors
 ; FILE REFERENCE: 1795/57155-A
 ; CURRENT APPLICATION NUMBER: US/09/255,368
 ; CURRENT FILING DATE: 1999-02-22
 ; EARLIER APPLICATION NUMBER: 09/161,113
 ; EARLIER FILING DATE: 1998-09-25
 ; NUMBER OF SEQ ID NOS: 42
 ; SOFTWARE: Patentin Ver. 2.0 - beta
 ; SEQ ID NO 2
 ; LENGTH: 432
 ; TYPE: PRT
 ; ORGANISM: Rattus norvegicus
 US-09-255-368-2

Query Match 2.4%; Score 8; DB 3; Length 432;
 Best Local Similarity 100.0%; Pred. No. 27;
 Matches 8; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 280 SSANPIIV 287
 |||||
 Db 324 SSANPIIV 331

Search completed: October 27, 2004, 08:52:22
 Job time : 42 secs

This Page Blank (uspto)

GenCore version 5.1.6
Copyright (c) 1993 - 2004 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: October 27, 2004, 08:51:45 / Search time 132 Seconds
(without alignments)
827.733 Million cell updates/sec

Title: US-09-867-570-2

Sequence: 1 MESKSSWVIRLGLSDSTI.....EGGGWLPQTLRLSGRLBQ 337

Scoring table: OLIGO
Gapop 60.0 , Gapept 60.0

Searched: 1370721 seqs, 324215800 residues

Word size : 0

Total number of hits satisfying chosen parameters: 1370721

Minimum DB seq length: 0

Maximum DB seq length: 200000000

Post-processing: Listing first 45 summaries

Database :

Published Applications AA:*
1: /cgn2_6/prodata/1/pubppa/US07_PUBCOMB.pep:*
2: /cgn2_6/prodata/1/pubppa/US06_PUBCOMB.pep:*
3: /cgn2_6/prodata/1/pubppa/US05_PUBCOMB.pep:*
4: /cgn2_6/prodata/1/pubppa/US04_PUBCOMB.pep:*
5: /cgn2_6/prodata/1/pubppa/US03_PUBCOMB.pep:*
6: /cgn2_6/prodata/1/pubppa/US02_PUBCOMB.pep:*
7: /cgn2_6/prodata/1/pubppa/US01_PUBCOMB.pep:*
8: /cgn2_6/prodata/1/pubppa/US00_PUBCOMB.pep:*
9: /cgn2_6/prodata/1/pubppa/US09_PUBCOMB.pep:*
10: /cgn2_6/prodata/1/pubppa/US08_PUBCOMB.pep:*
11: /cgn2_6/prodata/1/pubppa/US07_PUBCOMB.pep:*
12: /cgn2_6/prodata/1/pubppa/US06_PUBCOMB.pep:*
13: /cgn2_6/prodata/1/pubppa/US05_PUBCOMB.pep:*
14: /cgn2_6/prodata/1/pubppa/US04_PUBCOMB.pep:*
15: /cgn2_6/prodata/1/pubppa/US03_PUBCOMB.pep:*
16: /cgn2_6/prodata/1/pubppa/US02_PUBCOMB.pep:*
17: /cgn2_6/prodata/1/pubppa/US01_PUBCOMB.pep:*
18: /cgn2_6/prodata/1/pubppa/US00_PUBCOMB.pep:*
19: /cgn2_6/prodata/1/pubppa/US09_PUBCOMB.pep:*
20: /cgn2_6/prodata/1/pubppa/US08_PUBCOMB.pep:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	337	100.0	337	11	US-09-867-570-2
2	332	95.5	332	9	US-09-995-225-20
3	322	95.5	322	10	US-09-995-225-20
4	322	95.5	322	14	US-10-183-116-31
5	322	95.5	322	14	US-10-225-567A-674
6	322	95.5	322	15	US-10-072-012-529
7	322	95.5	322	15	US-10-072-012-534
8	322	95.5	322	14	US-10-219-834-79
9	287	85.2	322	14	US-10-401-397A-2
10	277	82.2	322	14	US-10-391-074-2
11	224	66.6	322	15	US-10-292-798-1274
12	221	65.6	322	15	US-10-072-012-530
13	221	65.6	322	15	US-10-072-012-535

14	181	53.7	302	14	US-10-237-467-10	Sequence 10, Appl
15	131	38.9	197	14	US-10-017-161-1600	Sequence 1600, Ap
16	76	22.6	76	14	US-10-219-834-174	Sequence 174, App
17	76	22.6	328	14	US-10-219-834-20	Sequence 20, Appl
18	50	14.8	319	15	US-10-072-012-174	Sequence 174, App
19	50	14.8	322	14	US-10-183-116-16	Sequence 16, Appl
20	50	14.8	322	14	US-10-079-384-4	Sequence 4, Appl1
21	50	14.8	322	14	US-10-017-161-1056	Sequence 1056, Ap
22	50	14.8	322	14	US-10-240-998-4	Sequence 4, Appl1
23	50	14.8	322	14	US-10-321-807-20	Sequence 20, Appl
24	50	14.8	322	14	US-10-237-467-12	Sequence 12, Appl
25	50	14.8	322	14	US-10-292-798-898	Sequence 898, App
26	50	14.8	322	15	US-10-016-248-81	Sequence 81, Appl
27	50	14.8	322	15	US-10-072-012-172	Sequence 172, App
28	50	14.8	322	15	US-10-072-012-527	Sequence 527, App
29	50	14.8	322	15	US-10-072-012-533	Sequence 533, App
30	50	14.8	322	15	US-10-343-650A-44	Sequence 44, Appl
31	50	14.8	322	16	US-10-321-807-20	Sequence 20, Appl
32	50	14.8	322	16	US-10-314-048A-20	Sequence 20, Appl
33	50	14.8	1589	15	US-10-072-012-528	Sequence 528, App
34	50	14.8	1589	15	US-10-072-012-532	Sequence 532, App
35	41	12.2	176	14	US-10-116-252-11	Sequence 11, Appl
36	41	12.2	322	9	US-09-995-225-18	Sequence 18, Appl
37	41	12.2	322	10	US-09-995-225-18	Sequence 18, Appl
38	41	12.2	322	14	US-10-183-116-33	Sequence 33, Appl
39	41	12.2	322	14	US-10-225-567A-689	Sequence 689, Appl
40	41	12.2	322	14	US-10-237-467-4	Sequence 4, Appl1
41	41	12.2	322	14	US-10-292-798-1042	Sequence 1042, Ap
42	41	12.2	322	15	US-10-016-248-84	Sequence 84, Appl
43	41	12.2	322	15	US-10-072-012-531	Sequence 531, App
44	41	12.2	322	15	US-10-072-012-536	Sequence 536, App
45	30	8.9	319	15	US-10-072-012-170	Sequence 170, App

ALIGNMENTS

RESULT 1
US-09-867-570-2
Sequence 2, Application US/09867570
Publication No. US20040076951A1
GENERAL INFORMATION:
APPLICANT: WEI, Ming-Hui et al.
TITLE OF INVENTION: ISOLATED HUMAN G-PROTEIN COUPLED RECEPTORS, NUCLEIC ACID MOLECULES ENCODING HUMAN GPCR
TITLE OF INVENTION: PROTEINS, AND USES THEREOF
FILE REFERENCE: CL000900-CIP
CURRENT APPLICATION NUMBER: US/09/867,570
CURRENT FILING DATE: 2001-05-31
PRIOR APPLICATION NUMBER: 09/695,045
PRIOR FILING DATE: 2000-10-25
NUMBER OF SEQ ID NOS: 4
SOFTWARE: FastSeq for Windows Version 4.0
SEQ ID NO 2
LENGTH: 337
TYPE: PRT
ORGANISM: Human
US-09-867-570-2

Query Match 100.0%; Score 337; DB 11; Length 337;
Best Local Similarity 100.0%; Pred. No. 4,1e-300;
Matches 337; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 1 MESKSSWVIRLGLSDSTIPIVGTETIPNGEERPCVQOTISFTGLTIVSVALTGN 60
DB 1 MESKSSWVIRLGLSDSTIPIVGTETIPNGEERPCVQOTISFTGLTIVSVALTGN 60
QY 61 AAVLMLGCMRRNNAVSIYILNVAADFLPLSGHIIICSPRLINIRIPISKIISPWTFF 120
DB 61 AAVLMLGCMRRNNAVSIYILNVAADFLPLSGHIIICSPRLINIRIPISKIISPWTFF 120
QY 121 YFTGLSTASISTERCISIIPIWYHQRDRYLSVWCVLMLSLRSLILEWFCDFLF 180
DB 121 YFTGLSTASISTERCISIIPIWYHQRDRYLSVWCVLMLSLRSLILEWFCDFLF 180

Db 121 YFGLSMLSAISTERCISILMPWYHCRPRYSSVWCVLWALSLRSILEMFCDFLE 180
Qy 181 SGADSWCETSDFTITAMVFLCVVLICGSSLVLVIRILCGSRKMPLTRLYTITLVVLF 240
Db 181 SGADSWCETSDFTITAMVFLCVVLICGSSLVLVIRILCGSRKMPLTRLYTITLVVLF 240
Qy 241 LCGLPFGIOWALFSRHLDMKVLFCVHLVSIPLSALNSSANPIIYFVGSFRORONRQ 300
Db 241 LCGLPFGIOWALFSRHLDMKVLFCVHLVSIPLSALNSSANPIIYFVGSFRORONRQ 300
Qy 301 NLKVLQALODTPEVDEGGGMLPOETLELSSGRLEQ 337
Db 301 NLKVLQALODTPEVDEGGGMLPOETLELSSGRLEQ 337

RESULT 2

US-09-995-225-20
Sequence 20, Application US/09995225
Publication No. US20020193584A1
GENERAL INFORMATION:
APPLICANT: Chen, Ruoping
APPLICANT: Chu, Zhi Liang
APPLICANT: Dang, Huong T.
APPLICANT: Lowitz, Kevin P.
APPLICANT: Priddy, Cameron
TITLE OF INVENTION: Endogenous And No. US20020193584A1-Endogenous Versions of Human G
FILE REFERENCE: AREN-0308
CURRENT APPLICATION NUMBER: US/09/995,225
CURRENT FILING DATE: 2001-11-26
PRIOR APPLICATION NUMBER: 09/170,496
PRIOR FILING DATE: 1998-10-13
PRIOR APPLICATION NUMBER: PCT/US99/23938
PRIOR FILING DATE: 1998-10-13
PRIOR APPLICATION NUMBER: 60/253,404
PRIOR FILING DATE: 2000-11-27
PRIOR APPLICATION NUMBER: 60/255,366
PRIOR FILING DATE: 2000-12-12
PRIOR APPLICATION NUMBER: 60/270,286
PRIOR FILING DATE: 2001-02-20
PRIOR APPLICATION NUMBER: 60/282,365
PRIOR FILING DATE: 2001-04-06
PRIOR APPLICATION NUMBER: 60/270,266
PRIOR FILING DATE: 2001-02-20
PRIOR APPLICATION NUMBER: 60/282,032
PRIOR FILING DATE: 2001-04-06
PRIOR APPLICATION NUMBER: 60/282,358
PRIOR FILING DATE: 2001-04-06
PRIOR APPLICATION NUMBER: 60/282,356
PRIOR FILING DATE: 2001-04-06
PRIOR APPLICATION NUMBER: 60/290,917
PRIOR FILING DATE: 2001-05-14
PRIOR APPLICATION NUMBER: 60/309,208
PRIOR FILING DATE: 2001-07-31
NUMBER OF SEQ ID NOS: 67
SOFTWARE: PatentIn version 3.1
SEQ ID NO 20
LENGTH: 322
TYPE: PRT
ORGANISM: Artificial Sequence
FEATURE:
OTHER INFORMATION: No. US20020193584A1 Sequence
US-09-995-225-20

Query Match 95.5%; Score 322; DB 9; Length 322;
Best Local Similarity 100.0%; Pred. No. 2.2e-286;
Matches 322; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 16 MDSTIPVLTGTELTPTINGREETPCYKOTLSFTGLTCTIVSLVATGNAVLMILGCMRRNA 75
Db 1 MDSTIPVLTGTELTPTINGREETPCYKOTLSFTGLTCTIVSLVATGNAVLMILGCMRRNA 60
Qy 76 VSIYILNLVADFLFLSGHIIICSPRLINIRHPISKILSPVMTFPYFGLSMLSAISTER 135

Db 61 VSIYILNLVADFLFLSGHIIICSPRLINIRHPISKILSPVMTFPYFGLSMLSAISTER 120
Qy 136 CISTIMPIWYHCRPRYSSVWCVLWALSLRSILEMFCDFLFSGADSWCETSDFT 195
Db 121 CISTIMPIWYHCRPRYSSVWCVLWALSLRSILEMFCDFLFSGADSWCETSDFT 180
Qy 196 IAMVFLCVVLICGSSLVLVIRILCGSRKMPLTRLYTITLVVFLCGLPFGIOWALFS 255
Db 181 IAMVFLCVVLICGSSLVLVIRILCGSRKMPLTRLYTITLVVFLCGLPFGIOWALFS 240
Qy 256 RHLDMKVLFCVHLVSIPLSALNSSANPIIYFVGSFRORONRQNLKVLQALODTPE 315
Db 241 RHLDMKVLFCVHLVSIPLSALNSSANPIIYFVGSFRORONRQNLKVLQALODTPE 300
Qy 316 VDEGGGMLPOETLELSSGRLEQ 337
Db 301 VDEGGGMLPOETLELSSGRLEQ 322

RESULT 3

US-09-995-225-20
Sequence 20, Application US/09995225
Publication No. US20030139588A9
GENERAL INFORMATION:
APPLICANT: Chen, Ruoping
APPLICANT: Chu, Zhi Liang
APPLICANT: Dang, Huong T.
APPLICANT: Lowitz, Kevin P.
APPLICANT: Priddy, Cameron
TITLE OF INVENTION: Endogenous And No. US20030139588A9-Endogenous Versions of Human G
FILE REFERENCE: AREN-0308
CURRENT APPLICATION NUMBER: US/09/995,225
CURRENT FILING DATE: 2001-11-26
PRIOR APPLICATION NUMBER: 09/170,496
PRIOR FILING DATE: 1998-10-13
PRIOR APPLICATION NUMBER: PCT/US99/23938
PRIOR FILING DATE: 1998-10-13
PRIOR APPLICATION NUMBER: 60/253,404
PRIOR FILING DATE: 2000-11-27
PRIOR APPLICATION NUMBER: 60/255,366
PRIOR FILING DATE: 2000-12-12
PRIOR APPLICATION NUMBER: 60/270,286
PRIOR FILING DATE: 2001-02-20
PRIOR APPLICATION NUMBER: 60/282,365
PRIOR FILING DATE: 2001-04-06
PRIOR APPLICATION NUMBER: 60/270,266
PRIOR FILING DATE: 2001-02-20
PRIOR APPLICATION NUMBER: 60/282,032
PRIOR FILING DATE: 2001-04-06
PRIOR APPLICATION NUMBER: 60/282,358
PRIOR FILING DATE: 2001-04-06
PRIOR APPLICATION NUMBER: 60/282,356
PRIOR FILING DATE: 2001-04-06
PRIOR APPLICATION NUMBER: 60/290,917
PRIOR FILING DATE: 2001-05-14
PRIOR APPLICATION NUMBER: 60/309,208
PRIOR FILING DATE: 2001-07-31
NUMBER OF SEQ ID NOS: 67
SOFTWARE: PatentIn version 3.1
SEQ ID NO 20
LENGTH: 322
TYPE: PRT
ORGANISM: Artificial Sequence
FEATURE:
OTHER INFORMATION: No. US20030139588A9 Sequence
US-09-995-225-20

Query Match 95.5%; Score 322; DB 10; Length 322;
Best Local Similarity 100.0%; Pred. No. 2.2e-286;
Matches 322; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 16 MDSTIPVLTGTELPINRGEETPCYKOTLSFTGLTCTIVSLVATGNVAVLMLGCRMRNA 75
Db 1 MDSTIPVLTGTELPINRGEETPCYKOTLSFTGLTCTIVSLVATGNVAVLMLGCRMRNA 60
Qy 76 VSIYIILNLVAADFLPSGHIICSPRLINIRHPISKILSPVMTFPYFGLSMLSAISTER 135
Db 61 VSIYIILNLVAADFLPSGHIICSPRLINIRHPISKILSPVMTFPYFGLSMLSAISTER 120
Qy 136 CLSILMPWVHCRRPRLSSVMCVLMLALSILSLIEMMCDPLFSGADSVWCETSDPIT 195
Db 121 CLSILMPWVHCRRPRLSSVMCVLMLALSILSLIEMMCDPLFSGADSVWCETSDPIT 180
Qy 196 IAMLVFLCVVLCSSSLVLRILGSRKMPRLRYTILTLVAVFLCGLPFGIQWALFS 255
Db 181 IAMLVFLCVVLCSSSLVLRILGSRKMPRLRYTILTLVAVFLCGLPFGIQWALFS 240
Qy 256 RIHDMKVLFCVHLVSIPLSALNSSANPIIYFVGSFRORONKLVLRALQDTPPE 315
Db 241 RIHDMKVLFCVHLVSIPLSALNSSANPIIYFVGSFRORONKLVLRALQDTPPE 300
Qy 316 VDEGGGWLPOETLESGSRLEQ 337
Db 301 VDEGGGWLPOETLESGSRLEQ 322

RESULT 8
US-10-219-834-79

Sequence 79, Application US/10219834
Publication No. US20030096751A1
GENERAL INFORMATION:
APPLICANT: Bristol-Myers Squibb Company
TITLE OF INVENTION: G-PROTEIN COUPLED RECEPTOR POLYNUCLEOTIDES AND METHODS OF USE THE
FILE REFERENCE: D0191 NP
CURRENT APPLICATION NUMBER: US/10/219,834
CURRENT FILING DATE: 2002-08-15
PRIOR APPLICATION NUMBER: US 60/313,658
PRIOR FILING DATE: 2001-08-20
PRIOR APPLICATION NUMBER: US 60/340,703
PRIOR FILING DATE: 2001-10-30
PRIOR APPLICATION NUMBER: US 60/318,675
PRIOR FILING DATE: 2001-09-12
PRIOR APPLICATION NUMBER: US 60/355,596
PRIOR FILING DATE: 2002-02-06
PRIOR APPLICATION NUMBER: US 60/333,417
PRIOR FILING DATE: 2001-11-26
PRIOR APPLICATION NUMBER: US 60/338,367
PRIOR FILING DATE: 2001-12-06
NUMBER OF SEQ ID NOS: 192
SOFTWARE: PatentIn version 3.1
SEQ ID NO 79
LENGTH: 314
TYPE: PRT
ORGANISM: Homo sapiens
US-10-219-834-79

Query Match 85.2%; Score 287; DB 14; Length 314;

Best Local Similarity 100.0%; Pred. No. 2,4e-254; Mismatches 0; Indels 0; Gaps 0;

Qy 16 MDSTIPVLTGTELPINRGEETPCYKOTLSFTGLTCTIVSLVATGNVAVLMLGCRMRNA 75
Db 1 MDSTIPVLTGTELPINRGEETPCYKOTLSFTGLTCTIVSLVATGNVAVLMLGCRMRNA 60
Qy 76 VSIYIILNLVAADFLPSGHIICSPRLINIRHPISKILSPVMTFPYFGLSMLSAISTER 135
Db 61 VSIYIILNLVAADFLPSGHIICSPRLINIRHPISKILSPVMTFPYFGLSMLSAISTER 120
Qy 136 CLSILMPWVHCRRPRLSSVMCVLMLALSILSLIEMMCDPLFSGADSVWCETSDPIT 195
Db 121 CLSILMPWVHCRRPRLSSVMCVLMLALSILSLIEMMCDPLFSGADSVWCETSDPIT 180
Qy 196 IAMLVFLCVVLCSSSLVLRILGSRKMPRLRYTILTLVAVFLCGLPFGIQWALFS 255

Db 181 IAMLVFLCVVLCSSSLVLRILGSRKMPRLRYTILTLVAVFLCGLPFGIQWALFS 240
Qy 256 RIHDMKVLFCVHLVSIPLSALNSSANPIIYFVGSFRORONKLVLRALQDTPPE 315
Db 241 RIHDMKVLFCVHLVSIPLSALNSSANPIIYFVGSFRORONKLVLRALQDTPPE 300

RESULT 9

US-10-401-397A-2

Sequence 2, Application US/10401397A
Publication No. US20030212001A1
GENERAL INFORMATION:
APPLICANT: Perit, Kriehna G.
APPLICANT: Moffett, Serge
APPLICANT: Abbar, Daniel
TITLE OF INVENTION: METHODS AND COMPOUNDS FOR PREVENTION AND TREATMENT OF ELEVATED
INTRACULAR PRESSURE AND RELATED CONDITIONS
FILE REFERENCE: 4518/1M674US1
CURRENT APPLICATION NUMBER: US/10/401,397A
CURRENT FILING DATE: 2003-03-27
PRIOR APPLICATION NUMBER: US 60/367,513
PRIOR FILING DATE: 2002-03-27
NUMBER OF SEQ ID NOS: 8
SOFTWARE: PatentIn version 3.1
SEQ ID NO 2
LENGTH: 322
TYPE: PRT
ORGANISM: Homo sapiens
US-10-401-397A-2

Query Match 82.2%; Score 277; DB 14; Length 322;

Best Local Similarity 100.0%; Pred. No. 3,6e-245; Mismatches 0; Indels 0; Gaps 0;

Qy 61 AVVLMGLGCRMRNAVSIYIILNLVAADFLPSGHIICSPRLINIRHPISKILSPVMTFP 120
Db 46 AVVLMGLGCRMRNAVSIYIILNLVAADFLPSGHIICSPRLINIRHPISKILSPVMTFP 105
Qy 121 YFGLSMLSAISTERCLSIIMPWVHCRRPRLSSVMCVLMLALSILSLIEMMCDPLF 180
Db 106 YFGLSMLSAISTERCLSIIMPWVHCRRPRLSSVMCVLMLALSILSLIEMMCDPLF 165
Qy 181 SGADSVWCETSDPITIAMVFLCVVLCSSSLVLRILGSRKMPRLRYTILTLVAVFL 240
Db 166 SGADSVWCETSDPITIAMVFLCVVLCSSSLVLRILGSRKMPRLRYTILTLVAVFL 225
Qy 241 LILGDPFGIQWALFSRIHDMKVLFCVHLVSIPLSALNSSANPIIYFVGSFRORONK 300
Db 226 LILGDPFGIQWALFSRIHDMKVLFCVHLVSIPLSALNSSANPIIYFVGSFRORONK 285
Qy 301 NLKLVLRALQDTPPEVDEGGGWLPOETLESGSRLEQ 337
Db 286 NLKLVLRALQDTPPEVDEGGGWLPOETLESGSRLEQ 322

RESULT 10

US-10-391-074-2

Sequence 2, Application US/10391074
Publication No. US20040038345A1
GENERAL INFORMATION:
APPLICANT: Nehls, Michael
APPLICANT: Matlier, Frank
TITLE OF INVENTION: No. US20040038345A1 Human Seven-Transmembrane Receptors
FILE REFERENCE: 7705,0008-00-000
CURRENT APPLICATION NUMBER: US/10/391,074
CURRENT FILING DATE: 2003-03-17
NUMBER OF SEQ ID NOS: 4
SOFTWARE: FastSeq for Windows Version 3.0.
SEQ ID NO 2
LENGTH: 322
TYPE: PRT
ORGANISM: Homo sapien

US-10-391-074-2

Query Match 66.5%; Score 224; DB 15; Length 322;
Best Local Similarity 100.0%; Pred. No. 1.3e-196;
Matches 224; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 98 SPLRLINIRHAPISKLSPVMTPEYPIGLSMISAISTERCLSLMPIMWYHCRPRYLSSVM 157
DB 83 SPLRLINIRHAPISKLSPVMTPEYPIGLSMISAISTERCLSLMPIMWYHCRPRYLSSVM 142
QY 158 CVLLMALSLSLSLLEMMFCDFLFGSADSVWCETSPFITAMVPLFCVVLGSSVLVARI 217
DB 143 CVLLMALSLSLSLLEMMFCDFLFGSADSVWCETSPFITAMVPLFCVVLGSSVLVARI 202
QY 218 LCGSRKMPLTRLYVTLLTVLVLFCGLPFGIOMLFSRIHDMKVLFCGHVLSIFLSA 277
DB 203 LCGSRKMPLTRLYVTLLTVLVLFCGLPFGIOMLFSRIHDMKVLFCGHVLSIFLSA 262
QY 278 LNSSANPIIYFVGSFRORONRQNLKVLQALQDTPVEDEGGG 321
DB 263 LNSSANPIIYFVGSFRORONRQNLKVLQALQDTPVEDEGGG 306

RESULT 11

US-10-292-798-1274
; Sequence 1274, Application US/10292798
; Publication No. US20030235633A1
; GENERAL INFORMATION:
; APPLICANT: SUMA, MAKIKO
; APPLICANT: ASAI, KIYOSHI
; APPLICANT: AKIYAMA, YUTAKA
; APPLICANT: ABRATANI, HIROYUKI
; TITLE OF INVENTION: GUANOSINE TRIPHOSPHATE-BINDING PROTEIN COUPLED RECEPTORS
; FILE REFERENCE: 084335/166
; CURRENT APPLICATION NUMBER: US/10/292,798
; CURRENT FILING DATE: 2002-11-13
; PRIOR APPLICATION NUMBER: 10/017,161
; PRIOR FILING DATE: 2001-12-18
; PRIOR APPLICATION NUMBER: JP 2001-246789
; PRIOR FILING DATE: 2001-06-18
; NUMBER OF SEQ ID NOS: 2070
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 1274
; LENGTH: 322
; TYPE: PRT
; ORGANISM: Homo sapiens
US-10-292-798-1274

Query Match 65.6%; Score 221; DB 14; Length 322;
Best Local Similarity 99.7%; Pred. No. 7.2e-194;
Matches 321; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 16 MDSTIPVIGTELTPINGREBTPCYKOTLSFTGLCTIVSLVATLGNVAVLMLGCMRRNA 75
DB 1 MDSTIPVIGTELTPINGREBTPCYKOTLSFTGLCTIVSLVATLGNVAVLMLGCMRRNA 60
QY 76 VSIYILNVADFLFLSGHIIICSPRLINIRHAPISKLSPVMTPEYPIGLSMISAISTER 135
DB 61 VSIYILNVADFLFLSGHIIICSPRLINIRHAPISKLSPVMTPEYPIGLSMISAISTER 120
QY 136 CLSLIMPIMWYHCRPRYLSSVMCVLLMALSLSLLEMMFCDFLFGSADSVWCETSPFIT 195
DB 121 CLSLIMPIMWYHCRPRYLSSVMCVLLMALSLSLLEMMFCDFLFGSADSVWCETSPFIT 180
QY 196 IAMVPLFCVVLGSSVLVARI LCGSRKMPLTRLYVTLLTVLVLFCGLPFGIOMALFS 255
DB 181 IAMVPLFCVVLGSSVLVARI LCGSRKMPLTRLYVTLLTVLVLFCGLPFGIOMALFS 240
QY 256 RIHDMKVLFCGHVLSIFLSALNSSANPIIYFVGSFRORONRQNLKVLQALQDTPVE 315
DB 241 RIHDMKVLFCGHVLSIFLSALNSSANPIIYFVGSFRORONRQNLKVLQALQDTPVE 300
QY 316 VDEGGWLPQETLRLSGSRLBQ 337

DB 301 VDEGGWLPQETLRLSGSRLBQ 322

RESULT 12

US-10-072-012-530
; Sequence 530, Application US/10072012
; Publication No. US2004003493A1
; GENERAL INFORMATION:
; APPLICANT: Tchernev, Velizar
; APPLICANT: Spytek, Kimberly
; APPLICANT: Zernusen, Bryan
; APPLICANT: Paturajan, Meera
; APPLICANT: Shinkets, Richard
; APPLICANT: Li, Li
; APPLICANT: Gangoli, Bsha
; APPLICANT: Padigaru, Muralidhara
; APPLICANT: Anderson, David W.
; APPLICANT: Raetelli, Luca
; APPLICANT: Miller, Charles E.
; APPLICANT: Gerlach, Valerie
; APPLICANT: Taupier Jr, Raymond J.
; APPLICANT: Gusev, Vladimir Y.
; APPLICANT: Coleman, Steven D.
; APPLICANT: Wolenc, Adam R.
; APPLICANT: Pena, Carol E. A
; APPLICANT: Furtak, Katarzyna
; APPLICANT: Groesse, William M.
; APPLICANT: Aleobrook II, John P.
; APPLICANT: Lepley, Denise M.
; APPLICANT: Rieger, Daniel K.
; APPLICANT: Burgess, Catherine E.
; TITLE OF INVENTION: Proteins and Nucleic Acids Encoding Same
; FILE REFERENCE: 21402-258
; CURRENT APPLICATION NUMBER: US/10/072,012
; CURRENT FILING DATE: 2002-01-31
; PRIOR APPLICATION NUMBER: 60/265,102
; PRIOR FILING DATE: 2001-01-30
; PRIOR APPLICATION NUMBER: 60/265,514
; PRIOR FILING DATE: 2001-01-31
; PRIOR APPLICATION NUMBER: 60/265,517
; PRIOR FILING DATE: 2001-01-31
; PRIOR APPLICATION NUMBER: 60/265,412
; PRIOR FILING DATE: 2001-01-31
; PRIOR APPLICATION NUMBER: 60/265,395
; PRIOR FILING DATE: 2001-01-31
; PRIOR APPLICATION NUMBER: 60/266,406
; PRIOR FILING DATE: 2001-02-02
; PRIOR APPLICATION NUMBER: 60/266,767
; PRIOR FILING DATE: 2001-02-05
; PRIOR APPLICATION NUMBER: 60/267,057
; PRIOR FILING DATE: 2001-02-07
; PRIOR APPLICATION NUMBER: 60/266,975
; PRIOR FILING DATE: 2001-02-07
; PRIOR APPLICATION NUMBER: 60/267,459
; PRIOR FILING DATE: 2001-02-08
; Remaining Prior Application data removed - See File Wrapper or PALM.
; NUMBER OF SEQ ID NOS: 1391
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 530
; LENGTH: 322
; TYPE: PRT
; ORGANISM: Homo sapiens
US-10-072-012-530

Query Match 65.6%; Score 221; DB 15; Length 322;
Best Local Similarity 99.7%; Pred. No. 7.2e-194;
Matches 321; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 16 MDSTIPVIGTELTPINGREBTPCYKOTLSFTGLCTIVSLVATLGNVAVLMLGCMRRNA 75
DB 1 MDSTIPVIGTELTPINGREBTPCYKOTLSFTGLCTIVSLVATLGNVAVLMLGCMRRNA 60

QY 76 VSIIYIINLVADPFLSGHIIICSPRLINIRHPISKILSPVMTFPYFIGLSMLSAISTER 135
DB 61 VSIIYIINLVADPFLSGHIIICSPRLINIRHPISKILSPVMTFPYFIGLSMLSAISTER 120
QY 136 CUSILMPYWHGCRPRYLSVWCVLMAISLRSIEMFCDPLFGADSVWCETSDFT 195
DB 121 CUSILMPYWHGCRPRYLSVWCVLMAISLRSIEMFCDPLFGADSVWCETSDFT 180
QY 196 IAMLVFLCVLGGSSVLVLRILICGSRKMPLTRLYTITLTVLVPFLCGLPFGIOWALFS 255
DB 181 IAMLVFLCVLGGSSVLVLRILICGSRKMPLTRLYTITLTVLVPFLCGLPFGIOWALFS 240
QY 256 RIHLDMKVLFGCHLVISIFLSALNSSANPIYFVGSFRORONRQNLKVLQALQDTPR 315
DB 241 RIHLDMKVLFGCHLVISIFLSALNSSANPIYFVGSFRORONRQNLKVLQALQDTPR 300
QY 316 VDEGGMLPQETTELGSRLQ 337
DB 301 VDEGGMLPQETTELGSRLQ 322

RESULT 13
US-10-072-012-535
Sequence 535, Application US/10072012
Publication No. US20040033493A1
GENERAL INFORMATION:
APPLICANT: Tchermnev, Velizar
APPLICANT: Spytek, Kimberly
APPLICANT: Zehrusen, Bryan
APPLICANT: Patunajan, Meera
APPLICANT: Shimkete, Richard
APPLICANT: Li, Li
APPLICANT: Gangolli, Beha
APPLICANT: Padigaru, Muralidhara
APPLICANT: Anderson, David W.
APPLICANT: Rastelli, Luca
APPLICANT: Miller, Charles E.
APPLICANT: Gerlach, Valerie
APPLICANT: Taupier Jr, Raymond J.
APPLICANT: Gusev, Vladimyr Y.
APPLICANT: Coleman, Steven D.
APPLICANT: Wolenc, Adam R.
APPLICANT: Pena, Carol E. A
APPLICANT: Furtak, Katarzyna
APPLICANT: Grosee, William M.
APPLICANT: Alsodrook II, John P.
APPLICANT: Lepley, Denise M.
APPLICANT: Rieger, Daniel K.
APPLICANT: Burgess, Catherine B.
TITLE OF INVENTION: Proteins and Nucleic Acids Encoding Same
FILE REFERENCE: 21402-258
CURRENT APPLICATION NUMBER: US/10/072,012
CURRENT FILING DATE: 2002-01-31
PRIOR APPLICATION NUMBER: 60/265,102
PRIOR FILING DATE: 2001-01-30
PRIOR APPLICATION NUMBER: 60/265,514
PRIOR FILING DATE: 2001-01-31
PRIOR APPLICATION NUMBER: 60/265,517
PRIOR FILING DATE: 2001-01-31
PRIOR APPLICATION NUMBER: 60/265,412
PRIOR FILING DATE: 2001-01-31
PRIOR APPLICATION NUMBER: 60/265,395
PRIOR FILING DATE: 2001-01-31
PRIOR APPLICATION NUMBER: 60/266,406
PRIOR FILING DATE: 2001-02-02
PRIOR APPLICATION NUMBER: 60/266,767
PRIOR FILING DATE: 2001-02-05
PRIOR APPLICATION NUMBER: 60/267,057
PRIOR FILING DATE: 2001-02-07
PRIOR APPLICATION NUMBER: 60/266,975
PRIOR FILING DATE: 2001-02-07
PRIOR APPLICATION NUMBER: 60/267,459
PRIOR FILING DATE: 2001-02-08

Remaining Prior Application data removed - See File Wrapper or PALM.
NUMBER OF SEQ ID NOS: 1391
SOFTWARE: PatentIn Ver. 2.1
SEQ ID NO 535
LENGTH: 322
TYPE: PRT
ORGANISM: Homo sapiens
US-10-072-012-535

Query Match 65.7%; Score 221; DB 15; Length 322;
Best Local Similarity 99.7%; Pred. No. 7, 2e-194;
Matches 321; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 16 MDSTIPVLGTETLPINGREETPCYKQTLSTFTGLTCIVSVALTGNVAVMLGCRNRNA 75
DB 1 MDSTIPVLGTETLPINGREETPCYKQTLSTFTGLTCIVSVALTGNVAVMLGCRNRNA 60
QY 76 VSIIYIINLVADPFLSGHIIICSPRLINIRHPISKILSPVMTFPYFIGLSMLSAISTER 135
DB 61 VSIIYIINLVADPFLSGHIIICSPRLINIRHPISKILSPVMTFPYFIGLSMLSAISTER 120
QY 136 CUSILMPYWHGCRPRYLSVWCVLMAISLRSIEMFCDPLFGADSVWCETSDFT 195
DB 121 CUSILMPYWHGCRPRYLSVWCVLMAISLRSIEMFCDPLFGADSVWCETSDFT 180
QY 196 IAMLVFLCVLGGSSVLVLRILICGSRKMPLTRLYTITLTVLVPFLCGLPFGIOWALFS 255
DB 181 IAMLVFLCVLGGSSVLVLRILICGSRKMPLTRLYTITLTVLVPFLCGLPFGIOWALFS 240
QY 256 RIHLDMKVLFGCHLVISIFLSALNSSANPIYFVGSFRORONRQNLKVLQALQDTPR 315
DB 241 RIHLDMKVLFGCHLVISIFLSALNSSANPIYFVGSFRORONRQNLKVLQALQDTPR 300
QY 316 VDEGGMLPQETTELGSRLQ 337
DB 301 VDEGGMLPQETTELGSRLQ 322

RESULT 14
US-10-237-467-10
Sequence 10, Application US/10237467
Publication No. US20030186324A1
GENERAL INFORMATION:
APPLICANT: Liao, Jiayu
APPLICANT: Gray, Nathanael S.
APPLICANT: Caldwell, Jeremy C.
APPLICANT: Schultz, Peter G.
APPLICANT: IRM LLC
TITLE OF INVENTION: Sensory Neuron Receptors
FILE REFERENCE: 021288-001300US
CURRENT APPLICATION NUMBER: US/10/237,467
CURRENT FILING DATE: 2003-01-14
PRIOR APPLICATION NUMBER: US 60/317,879
PRIOR FILING DATE: 2001-09-07
NUMBER OF SEQ ID NOS: 18
SOFTWARE: PatentIn Ver. 2.1
SEQ ID NO 10
LENGTH: 302
TYPE: PRT
ORGANISM: Homo sapiens
FEATURE:
OTHER INFORMATION: dorsal root ganglia G-protein coupled receptor (GPCR)
OTHER INFORMATION: 7 (DRG) (NT009307)
US-10-237-467-10

Query Match 53.7%; Score 181; DB 14; Length 302;
Best Local Similarity 99.6%; Pred. No. 3e-157;
Matches 281; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 56 ALTGNAVITMLGCRNRNAVSIYIINLVADPFLSGHIIICSPRLINIRHPISKILSP 115
DB 21 ALTGNAVITMLGCRNRNAVSIYIINLVADPFLSGHIIICSPRLINIRHPISKILSP 80

Qy	116	WTFPYFIFGLSMJSAJSTERCLSIIMPIWCHCRPRJSSWCVLLMALSLSIIEEMF	175
Db	81	WTFPYFIFGLSMJSAJSTERCLSIIMPIWCHCRPRJSSWCVLLMALSLSIIEEMF	140
Qy	176	CDLFESGADSWMCETSDPTTIAMLVFLCVLLCGSSJVLVLVRIILCGSRKMPJRLYYTITL	235
Db	141	CDLFESGADSWMCETSDPTTIAMLVFLCVLLCGSSJVLVLVRIILCGSRKMPJRLYYTITL	200
Qy	236	TLVLVFLLCGLPRGICQWALFSRIHLDMKVLFCGHVLYSIFLSAANSSANPIYFVGSFRQ	295
Db	201	TLVLVFLLCGLPRGICQWALFSRIHLDMKVLFCGHVLYSIFLSAANSSANPIYFVGSFRQ	260
Qy	296	RQNRQNLKVLQRALQDTPEDVEGGMLPEQETELSGSRLEQ	337
Db	261	RQNRQNLKVLQRALQDTPEDVEGGMLPEQETELSGSRLEQ	302

RESULT 15
 US-10-017-161-1600
 Sequence 1600, Application US/10017161
 Publication No. US20030143668A1
 GENERAL INFORMATION:
 APPLICANT: SUWA, MAKIKO
 APPLICANT: ASAI, KIYOSHI
 APPLICANT: AKIYAMA, YUTAKA
 APPLICANT: ABRUTANI, HIROYUKI
 TITLE OF INVENTION: NOVEL G PROTEIN-COUPLED RECEPTORS
 FILE REFERENCE: 084335/0152
 CURRENT APPLICATION NUMBER: US/10/017,161
 CURRENT FILING DATE: 2002-11-18
 PRIOR APPLICATION NUMBER: JP 2001/246789
 PRIOR FILING DATE: 2001-06-18
 NUMBER OF SEQ ID NOS: 2430
 SOFTWARE: PatentIn Ver. 2.1
 SEQ ID NO 1600
 LENGTH: 197
 TYPE: PRY
 ORGANISM: Homo sapiens
 US-10-017-161-1600

	Query Match Similarity	38.9%;	Score 131;	DB 14;	Length 197;	
	Best Local Similarity	100.0%;	Pred. No. 1.3e-111;			
	Matches 131;	Conservative 0;	Mismatches 0;	Indels 0;	Gaps 0	
QY	16	MDSTIPVLGTETLPINGREETPCYKQTSIFGTGLTIVSLVALTGNAAVIMLLGCRNRRA	75			
Db	1	MDSTIPVLGTETLPINGREETPCYKQTSIFGTGLTIVSLVALTGNAAVIMLLGCRNRRA	60			
QY	76	VSITILNLAADFLPLSGHITCSPLRLNIRHPISKILSPVWTTPPFITGLSMLSAISTER	135			
Db	61	VSITILNLAADFLPLSGHITCSPLRLNIRHPISKILSPVWTTPPFITGLSMLSAISTER	120			
QY	136	CLSTILMPIMYH 146				
Db	121	CLSTILMPIMYH 131				

Search completed: October 27, 2004, 09:03:34
Job time : 133 secs

GenCore version 5.1.6
Copyright (c) 1993 - 2004 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: October 27, 2004, 08:43:20 ; Search time 39 Seconds
(without alignments)
831.412 Million cell updates/sec

Title: US-09-867-570-2

Sequence: 1 MESKSSWVIRLIGFLSMDSTI.....EGGGWLPGRTLRISGRLEQ 337

Scoring table: OLIGO
Gapop 60.0 , Gapext 60.0

Searched: 283416 seqs, 96216763 residues

Word size : 0

Total number of hits satisfying chosen parameters: 283416

Minimum DB seq length: 0

Maximum DB seq length: 200000000

Post-processing: Listing first 45 summaries

Database :

1: PIR.79:.*
2: PIR1:.*
3: PIR2:.*
4: PIR3:.*
5: PIR4:.*

Pred. No. is the number of results predicted by chance to have a
score greater than or equal to the score of the result being printed,
and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	14	4.2	378	2	A39485 transforming prote
2	10	3.0	435	2	A44308 Antho-Rfamide prec
3	8	2.4	111	2	S16099 viascotoxin - Europ
4	8	2.4	141	2	C81055 hypothetical prote
5	8	2.4	141	2	B81822 hypothetical prote
6	8	2.4	215	2	JS0037 somatotropin precu
7	8	2.4	215	2	JS0037 somatotropin - bul
8	8	2.4	230	2	I48685 mast cell proteina
9	8	2.4	276	2	A38654 mast cell proteina
10	8	2.4	411	2	T39481 hypothetical prote
11	8	2.4	429	2	F82679 conserved hypochet
12	8	2.4	455	2	G70089 hypothetical prote
13	8	2.4	541	2	AH2909 hypothetical prote
14	8	2.4	541	2	F97684 chitinam transport
15	8	2.4	984	2	T50309 hypothetical WD-re
16	7	2.1	49	1	LBHFC light-harvesting p
17	7	2.1	50	1	LBHFC light-harvesting p
18	7	2.1	71	2	B45874 ig alpha-1 chain C
19	7	2.1	125	2	C72579 ig alpha-2 chain C
20	7	2.1	160	2	AB1185 B. subtilis Yds p
21	7	2.1	163	2	H75203 probable biotin sy
22	7	2.1	174	2	B75114 probable NADH dehy
23	7	2.1	179	2	B70757 hypothetical prote
24	7	2.1	181	2	F69138 hypothetical prote
25	7	2.1	186	2	F90150 conserved hypochet
26	7	2.1	199	2	JU0052 puromycin N-acetyl
27	7	2.1	209	2	C70424 hypothetical prote
28	7	2.1	213	2	T27841 hypothetical prote

30	7	2.1	256	2	A11119 transcription regu
31	7	2.1	256	2	AD1480 transcription regu
32	7	2.1	260	2	AD1481 hypothetical prote
33	7	2.1	263	2	S48498 oxidoreductase hom
34	7	2.1	269	2	C71255 conserved hypochet
35	7	2.1	270	2	S16579 chitinase (BC 3.2.
36	7	2.1	276	2	C35270 DNA-damage repair
37	7	2.1	284	2	AH0415 probable membrane-
38	7	2.1	298	2	F82356 transcription acti
39	7	2.1	300	2	T26245 hypothetical prote
40	7	2.1	301	2	F86440 unknown protein [i
41	7	2.1	303	2	B96909 probable permease
42	7	2.1	313	2	AB2203 hypothetical prote
43	7	2.1	315	2	T49982 hypothetical prote
44	7	2.1	322	2	T24525 hypothetical prote
45	7	2.1	324	1	TVRTAS transforming prote

ALIGNMENTS

RESULT 1

A39485 transforming protein (mrg) - human
C:Species: Homo sapiens (man)
C:Date: 28-Feb-1992 #sequence_revision 17-Apr-1993 #text_change 09-Jul-2004
C:Accession: A39485
R:Monnot, C.; Weber, V.; Stinakre, J.; Bihoreau, C.; Teutsch, B.; Corvol, P.; Clauser, J.
Mol. Endocrinol. 5, 1477-1487, 1991
A:Title: Cloning and functional characterization of a novel mas-related gene, modulating
A:Reference number: A39485; MUID:92130997; PMID:1723144
A:Accession: A39485
A:Status: preliminary
A:Molecule type: DNA
A:Residues: 1-378 <MON>
C:Superfamily: mas transforming protein
C:Keywords: G.protein-coupled receptor; transmembrane protein

Query Match 4.2%; Score 14; DB 2; Length 378;
Best Local Similarity 100.0%; Pred. No. 6.2e-06;
Matches 14; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 279 NSSANPIYFVGS 292
Db 306 NSSANPIYFVGS 319

RESULT 2

A44308 Antho-Rfamide precursor - sea anemone (Anthopleura elegantissima)
C:Species: Anthopleura elegantissima
C:Date: 10-Jun-1993 #sequence_revision 18-Nov-1994 #text_change 09-Jul-2004
C:Accession: A44308
R:Schmutzler, C.; Darmer, D.; Diekhoff, D.; Grimmelikhuijzen, C.J.
J. Biol. Chem. 267, 22534-22541, 1992
A:Title: Identification of a novel type of processing sites in the precursor for the sea
A:Reference number: A44308; MUID:93054550; PMID:1429603
A:Accession: A44308
A:Status: preliminary
A:Molecule type: mRNA; protein
A:Residues: 1-435 <SCH>
A:Cross-references: UNIPROT:P10419; GB:M98269; NID:G155702; PIDN:AAA27738.1; PID:G155703
A>Note: sequence extracted from NCBI backbone (NCBIN:117102, NCBI:P.117104)
C:Keywords: neuropeptide

Query Match 3.0%; Score 10; DB 2; Length 435;
Best Local Similarity 100.0%; Pred. No. 0.085;
Matches 10; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 230 YVITLLTVLV 239
Db 6 YVITLLTVLV 15

```

RESULT 3
S16039
viscotoxin - European mistletoe
C:Species: Viscum album (European mistletoe)
C>Date: 19-Mar-1997 #sequence_revision 19-Mar-1997 #text_change 31-Oct-1997
C:Accession: S16039
R:Schrader, G.; Apel, K.
Eur. J. Biochem. 198, 549-553, 1991
A:Title: Isolation and characterization of cDNAs encoding viscotoxins of mistletoe (Visc
A:Reference number: S16039; MUID:91266334; PMID:1710983
A:Accession: S16039
A:Status: preliminary
A:Molecule type: mRNA
A:Residues: 1-111 <EUR>
C:Superfamily: viscotoxin

Query Match      2.4%; Score 8; DB 2; Length 111;
Best Local Similarity 100.0%; Pred. No. 2.9;
Matches 8; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY      208 GSSLVTLV 215
Db      6 GSSLVTLV 13

RESULT 4
C81055
hypothetical protein NMB1681 [imported] - Neisseria meningitidis (strain MCS8 serogroup
C:Species: Neisseria meningitidis
C>Date: 31-Mar-2000 #sequence_revision 31-Mar-2000 #text_change 09-Jul-2004
C:Accession: C81055
R:Tettein, H.; Saunders, N.J.; Heidelberg, J.; Jeffries, A.C.; Nelson, K.E.; Eisen, J.A
Hickey, E.R.; Haft, D.H.; Salzberg, S.L.; White, O.; Fleischmann, R.D.; Dougherty, B.A.;
ri, H.; Qin, H.; Vamathavan, J.; Gill, J.; Scariato, V.; Maignani, V.; Pizza, M.
Science 287, 1809-1815, 2000
A:Authors: Grandi, G.; Sun, L.; Smith, H.O.; Frazer, C.M.; Moxon, E.R.; Rappuoli, R.; Ve
A:Title: Complete genome sequence of Neisseria meningitidis serogroup B strain MCS8.
A:Reference number: A81000; MUID:20175755; PMID:10710307
A:Accession: C81055
A:Status: preliminary
A:Molecule type: DNA
A:Residues: 1-141 <TET>
A:Cross-references: UNIPROT:Q9JY98; GB:AE002518; GB:AE002098; NID:G7226928; PIDN:AAF4202
A:Experimental source: serogroup B, strain MCS8
C:Genetics:
A:Gene: NMB1681

Query Match      2.4%; Score 8; DB 2; Length 141;
Best Local Similarity 100.0%; Pred. No. 3.6;
Matches 8; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY      146 HCRPRRYL 153
Db      71 HCRPRRYL 78

RESULT 5
B81822
hypothetical protein NMA1940 [imported] - Neisseria meningitidis (strain Z2491 serogroup
C:Species: Neisseria meningitidis
C>Date: 05-May-2000 #sequence_revision 05-May-2000 #text_change 09-Jul-2004
C:Accession: B81822
R:Parhill, J.; Achtman, M.; James, K.D.; Bentley, S.D.; Churcher, C.; Klee, S.R.; Morel
; Holroyd, S.; Jørgen, K.; Leather, S.; Moule, S.; Mungall, K.; Quail, M.A.; Rajandream,
Nature 404, 502-506, 2000
A:Title: Complete DNA sequence of a serogroup A strain of Neisseria meningitidis Z2491.
A:Reference number: A81775; MUID:20222556; PMID:10761919
A:Accession: B81822
A:Status: preliminary
A:Molecule type: DNA
A:Residues: 1-141 <PAR>

```

```

A:Cross-references: UNIPROT:Q9UT80; GB:AL162757; GB:AL157959; NID:G7380371; PIDN:CA88516
A:Experimental source: serogroup A, strain Z2491
C:Genetics:
A:Gene: NMA1940

Query Match      2.4%; Score 8; DB 2; Length 141;
Best Local Similarity 100.0%; Pred. No. 3.6;
Matches 8; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY      146 HCRPRRYL 153
Db      71 HCRPRRYL 78

RESULT 6
JS0037
somatotropin precursor - bullfrog
N:Alternate names: growth hormone
C:Species: Rana catesbeiana (bullfrog)
C>Date: 31-Mar-1992 #sequence_revision 31-Mar-1992 #text_change 09-Jul-2004
C:Accession: JS0037; PS0310
R:Pan, F.M.; Chang, W.C.
Biochim. Biophys. Acta 950, 238-242, 1988
A:Title: Cloning and sequencing of bullfrog growth hormone complementary DNA.
A:Reference number: JS0037; MUID:88252154; PMID:3260110
A:Accession: JS0037
A:Molecule type: mRNA
A:Residues: 1-215 <PAN>
A:Cross-references: UNIPROT:P10813; GB:X12520; NID:G64259; PIDN:CAA31038.1; PID:G64260
A:Accession: PS0310
A:Molecule type: protein
A:Residues: 26-55 <PAZ>
C:Comment: This protein is synthesized and secreted by the anterior pituitary gland and
C:Superfamily: prolactin
C:Keywords: growth factor; hormone
F1-25/Domain: signal sequence #status predicted <SIG>
F126-215/Product: somatotropin #status experimental <MAT>

Query Match      2.4%; Score 8; DB 2; Length 215;
Best Local Similarity 100.0%; Pred. No. 5.1;
Matches 8; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY      208 GSSLVTLV 215
Db      6 GSSLVTLV 13

RESULT 7
IS1188
somatotropin - bullfrog
N:Alternate names: growth hormone
C:Species: Rana catesbeiana (bullfrog)
C>Date: 13-Sep-1996 #sequence_revision 13-Sep-1996 #text_change 09-Jul-2004
C:Accession: IS1188
R:Takehashi, N.; Kikuyama, S.; Gen, K.; Maruyama, O.; Kato, Y.
J. Mol. Endocrinol. 9, 283-289, 1992
A:Title: Cloning of a bullfrog growth hormone cDNA: expression of growth hormone mRNA in
A:Reference number: IS1188; MUID:93119453; PMID:1476615
A:Accession: IS1188
A:Status: preliminary; translated from GB/EMBL/DBJ
A:Molecule type: mRNA
A:Residues: 1-215 <TAK>
A:Cross-references: UNIPROT:P10813; GB:SS2027; NID:G262921; PIDN:AAB24792.1; PID:G262922
C:Superfamily: prolactin

Query Match      2.4%; Score 8; DB 2; Length 215;
Best Local Similarity 100.0%; Pred. No. 5.1;
Matches 8; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY      208 GSSLVTLV 215
Db      6 GSSLVTLV 13

```


RESULT 8
I48685
mact cell proteinase 6 (EC 3.4.21.-) precursor - mouse
C/Species: Mus musculus (house mouse)
C/Date: 02-Jul-1996 #sequence_revision 02-Jul-1996 #text_change 09-Jul-2004
C/Accession: I48685; S43172
R/Huang, R.; Hellman, L.
Immunogenetics 40, 397-414, 1994
A/Title: Genes for mact-cell serine protease and their molecular evolution.
A/Reference number: I48684; MUID:95048582; PMID:7959952
A/Accession: I48685
A/Status: preliminary; translated from GB/EMBL/DBJ
A/Molecule type: mRNA
A/Residues: 1-230 <RES>
A/Cross-references: UNIPROT:P21845; EMBL:X78542; NID:9468809; PIDN:CAA55288.1; PID:94688
C/Superfamily: trypsin; trypsin homology
C/Keywords: hydrolase; serine proteinase
F/22-230/Domain: trypsin homology #status atypical <TRY>

Query Match
Best Local Similarity 100.0%; Score 8; DB 2; Length 230;
Pred. No. 5.5; Mismatches 0; Indels 0; Gaps 0;
Matches 8; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 160 LLMALSL 167
Db 8 LLMALSL 15

RESULT 9
A38654
mact cell proteinase 6 (EC 3.4.21.-) precursor - mouse
C/Species: Mus musculus (house mouse)
C/Date: 21-Feb-1992 #sequence_revision 17-Feb-1994 #text_change 09-Jul-2004
C/Accession: A38654; B38654; D35646; I59478
R/Reynolds, D.S.; Gurley, D.S.; Austen, K.F.; Serafin, W.B.
J. Biol. Chem. 266, 3847-3853, 1991
A/Title: Cloning of the cDNA and gene of mouse mact cell proteinase-6. Transcription by p
A/Reference number: A38654; MUID:91139682; PMID:1995638
A/Accession: A38654
A/Molecule type: DNA
A/Residues: 1-276 <REV>
A/Cross-references: UNIPROT:P21845; GB:M57625; NID:9200506; PIDN:AAA39987.1; PID:9200507
A/Note: the authors translated the codon CGC for residue 24 as Ala, GAG for residue 37 as
Gly, GAG for residue 148 as Gly, GAG for residue 168 as Gly, and GAA for 185 as Gly
A/Accession: B38654
A/Molecule type: mRNA
A/Residues: 1-276 <RES>
A/Cross-references: GB:M57626; NID:9200508; PIDN:AAA39988.1; PID:9200509
R/Reynolds, D.S.; Stevens, R.L.; Lane, W.S.; Carr, M.H.; Austen, K.F.; Serafin, W.B.
Proc. Natl. Acad. Sci. U.S.A. 87, 3230-3234, 1990
A/Title: Different mouse mact cell populations express various combinations of at least
A/Reference number: A35646; MUID:90222202; PMID:2326280
A/Accession: D35646
A/Molecule type: protein
A/Residues: 32-54 <RES>
R/Huang, R.; Abrik, M.; Gohl, A.E.; Nilsson, G.; Aveskog, M.; Larsson, L.G.; Nilsson,
Scand. J. Immunol. 38, 359-367, 1993
A/Title: Expression of a mact cell trypsin in the human monocytic cell lines U-937 and
A/Reference number: I59478; MUID:94023807; PMID:8210998
A/Accession: I59478
A/Status: preliminary; translated from GB/EMBL/DBJ
A/Molecule type: mRNA
A/Residues: 1-276 <RES>
A/Cross-references: GB:U31853; NID:9473480; PIDN:AAA39725.1; PID:9473481
C/Genetics:
A/Gene: MMCP-6
A/Intons: 24/1; 79/2; 168/1; 222/3
C/Superfamily: trypsin; trypsin homology
C/Keywords: hydrolase; serine proteinase; zymogen
F/1-21/Domain: signal sequence #status predicted <SIG>
F/22-31/Domain: activation peptide #status predicted <ACT>
F/32-276/Product: mact cell proteinase 6 #status experimental <MAT>

F/32-268/Domain: trypsin homology <TRY>
F/75,122,225/Active site: His, Asp, Ser #status predicted

Query Match
Best Local Similarity 100.0%; Score 8; DB 2; Length 276;
Pred. No. 6.4; Mismatches 0; Indels 0; Gaps 0;
Matches 8; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 160 LLMALSL 167
Db 8 LLMALSL 15

RESULT 10
T39481
hypothetical protein SPBC15D4.05 - fission yeast (Schizosaccharomyces pombe)
C/Species: Schizosaccharomyces pombe
C/Date: 03-Dec-1999 #sequence_revision 03-Dec-1999 #text_change 09-Jul-2004
C/Accession: T39481
R/Lyne, M.; Rajandream, M.A.; Barrett, B.G.; Lucas, M.; Gallardin, C.
submitted to the EMBL Data Library, August 1997
A/Reference number: Z21858
A/Accession: T39481
A/Status: preliminary; translated from GB/EMBL/DBJ
A/Molecule type: DNA
A/Residues: 1-411 <LYN>
A/Cross-references: UNIPROT:O74310; EMBL:AL031349; PIDN:CAA20480.1; GSPDB:GN00067; SPDB:
A/Experimental source: strain 972h-; cosmid c15D4
C/Genetics:
A/Gene: SPDB:SPBC15D4.05
A/Map position: 2

Query Match
Best Local Similarity 100.0%; Score 8; DB 2; Length 411;
Pred. No. 9; Mismatches 0; Indels 0; Gaps 0;
Matches 8; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 165 SLRSLR 172
Db 76 SLRSLR 83

RESULT 11
F82679
conserved hypothetical protein XF1453 [imported] - Xylella fastidiosa (strain 9a5c)
C/Species: Xylella fastidiosa
C/Date: 18-Aug-2000 #sequence_revision 20-Aug-2000 #text_change 09-Jul-2004
C/Accession: F82679
R/anonymous, The Xylella fastidiosa Consortium of the Organization for Nucleotide Sequenc
Nature 406, 151-157, 2000
A/Title: The genome sequence of the plant pathogen Xylella fastidiosa.
A/Reference number: A82515; MUID:20365717; PMID:10910347
A/Note: for a complete list of authors see reference number A59328 below
A/Accession: F82679
A/Status: preliminary
A/Molecule type: DNA
A/Residues: 1-429 <SIM>
A/Cross-references: UNIPROT:Q9PDC6; GB:AE003975; GB:AE003849; NID:9106468; PIDN:AF8426;
A/Experimental source: strain 9a5c
R/Simpson, A.J.G.; Reineck, F.C.; Arruda, P.; Abreu, F.A.; Acencio, M.; Alvarenga, R.; Al
Birtone, M.R.S.; Bueno, M.R.P.; Camargo, A.A.; Camargo, L.E.A.; Carraro, D.M.; Carre, H
as-Neto, E.; Docena, C.; El-Dor, H.; Facincani, A.P.; Ferreira, A.J.S.
submitted to GenBank, June 2000
A/Authors: Ferreira, V.C.A.; Ferro, J.A.; Fraga, J.S.; Franca, S.C.; Franco, M.C.; Fromme
J.D.; Junqueira, M.L.; Kemper, E.L.; Kiteajima, U.P.; Krieger, J.E.; Kurame, E.B.; Laigre
Chado, M.A.; Madeira, A.M.B.N.; Madeira, H.M.F.; Martino, G.L.; Marques, M.V.; Martins, B
A/Authors: Martins, E.M.F.; Matsukuma, A.Y.; Menck, C.F.M.; Mireca, E.C.; Miyaki, C.Y.;
F.G.; Nunes, L.R.; Oliveira, M.A.; de Oliveira, M.C.; de Oliveira, R.C.; Palmieri, D.A
Rodrigues, V.; Rosa, A.J.; de M.; de Rosa Jr., V.E.; de Sa, R.G.; Santelli, R.V.; Sawasaki
A/Authors: da Silva, A.C.R.; da Silva, P.R.; da Silva, A.M.; Silva Jr., W.A.; da Silveira
M.; Tshak, M.H.; Vallada, H.; Van Sluys, M.A.; Verjovski-Almeida, S.; Vettore, A.L.; Z
A/Reference number: A59328
A/Contents: annotation
C/Genetics:
A/Gene: XF1453

C/Superfamily: Haemophilus influenzae conserved hypothetical protein HII590

Query Match 2.4%; Score 8; DB 2; Length 429;
Best Local Similarity 100.0%; Pred. No. 9.3;
Matches 8; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 306 LQALQDT 313
|||||
DB 155 LQALQDT 162

RESULT 12

G70089

hypothetical protein yycH - Bacillus subtilis

C/Species: Bacillus subtilis

C/Date: 05-Dec-1997 #sequence_revision 05-Dec-1997 #text_change 09-Jul-2004

C/Accession: G70089
R/Kunst, F.; Ogasawara, N.; Moszer, I.; Albertini, A.M.; Alloni, G.; Azevedo, V.; Berber, C.; Bron, S.; Brouillet, S.; Brusch, C.V.; Caldwell, B.; Capuano, V.; Carter, N.M.; Cho, A.; Ehrlich, S.D.; Emerson, P.T.; Eutlian, K.D.; Eyrington, J.; Fabret, C.; Ferrari, B.; Nature 390, 249-256, 1997

A/Authors: Foulger, D.; Fritz, C.; Fujita, M.; Fujita, Y.; Fuma, S.; Gallizi, A.; Galliech, J.; Harwood, C.R.; Henauf, A.; Hilbert, H.; Holsappel, S.; Hosono, S.; Hullo, M.F.; Koetter, P.; Koningsstein, G.; Krogh, S.; Kumano, M.; Kurita, K.; Lapidus, A.; Lardinois, A.; Authors: Lauber, J.; Lazarevic, V.; Lee, S.M.; Levine, A.; Liu, H.; Masuda, S.; Maueil, Y., M.; Ogawa, K.; Ogiwara, A.; Oudega, B.; Park, S.H.; Parro, V.; Pohl, T.M.; Portetelle, Rieger, M.; Rivolta, C.; Rocha, E.; Roche, B.; Rose, M.; Sadate, Y.; Sato, T.; Scanlon, A.; Authors: Schleich, S.; Schroeder, R.; Scoffone, F.; Sekiguchi, J.; Sekowska, A.; Serot, A.; Winters, P.; Wipat, A.; Yamamoto, H.; Yamane, K.; Yasunori, K.; Yata, K.; Yoshida, K.; Authors: Yoshikawa, H.F.; Zumbach, E.; Yoshikawa, H.; Danchin, A.

A/Title: The complete genome sequence of the Gram-positive bacterium Bacillus subtilis.

A/Reference number: A69580; MUID:98044033; PMID:9384377

A/Accession: G70089

A/Status: preliminary; nucleic acid sequence not shown; translation not shown

A/Molecule type: DNA

A/Residues: 1-455 <KUN>

A/Cross-references: UNIPROT:Q45613; GB:Z99124; GB:AL009126; NID:G2636442; PIDN:CAB16076.

A/Experimental source: strain 168

C/Genetics:

A/Gene: yycH

C/Superfamily: Bacillus subtilis hypothetical protein yycH

Query Match 2.4%; Score 8; DB 2; Length 455;
Best Local Similarity 100.0%; Pred. No. 9.8;
Matches 8; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 232 TILTLVLY 239
|||||
DB 8 TILTLVLY 15

RESULT 13

AH2909

hypothetical protein thip [imported] - Agrobacterium tumefaciens (strain C58, Dupont)

C/Species: Agrobacterium tumefaciens

C/Date: 11-Jan-2002 #sequence_revision 11-Jan-2002 #text_change 09-Jul-2004

C/Accession: AH2909

R/Mood, D.W.; Setubal, J.C.; Kaul, R.; Monks, D.; Chen, L.; Wood, G.E.; Chen, Y.; Woo, I.; Karp, P.; Romero, P.; Grant, C.; Guenther, D.; Kutayavin, T.; Levy, R.; Li, M.; McClellan, Science 294, 2317-2323, 2001

A/Authors: Yoo, H.; Tao, Y.; Biddle, P.; Jung, M.; Krespan, W.; Perry, M.; Gordon-Kamm, A.; Reference number: AB2577; MUID:21608550; PMID:11743193

A/Title: The Genome of the Natural Genetic Engineer Agrobacterium tumefaciens C58.

A/Status: preliminary

A/Molecule type: DNA

A/Residues: 1-541 <KUN>

A/Cross-references: UNIPROT:Q8UBV5; GB:AE008688; PIDN:AL4694.1; PID:g17741221; GSPDB:C

A/Experimental source: strain C58 (Dupont)

C/Genetics:

A/Gene: thip
A/Map position: circular chromosome

Query Match 2.4%; Score 8; DB 2; Length 541;
Best Local Similarity 100.0%; Pred. No. 11;
Matches 8; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 208 GSSVLVLY 215
|||||
DB 378 GSSVLVLY 385

RESULT 14

F97684

thiamin transport system permease protein thip [imported] - Agrobacterium tumefaciens (str

C/Species: Agrobacterium tumefaciens

C/Date: 30-Sep-2001 #sequence_revision 30-Sep-2001 #text_change 09-Jul-2004

C/Accession: F97684
R/Goodner, B.; Hinkle, G.; Gattung, S.; Miller, N.; Blanchard, M.; Qurollo, B.; Goldman, A.; Liu, F.; Wollam, C.; Allinger, M.; Dougherty, D.; Scott, C.; Lappas, C.; Markelz, B.; Science 294, 2323-2328, 2001

A/Title: Genome Sequence of the Plant Pathogen and Biotechnology Agent Agrobacterium tumefaciens

A/Reference number: A97359; MUID:21608551; PMID:11743194

A/Accession: F97684

A/Status: preliminary

A/Molecule type: DNA

A/Residues: 1-541 <KUN>

A/Cross-references: UNIPROT:Q8UBV5; GB:AE007869; PIDN:AK88431.1; PID:g15157926; GSPDB:GT

A/Genes: AGR_C_4917

A/Map position: circular chromosome

Query Match 2.4%; Score 8; DB 2; Length 541;
Best Local Similarity 100.0%; Pred. No. 11;
Matches 8; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 208 GSSVLVLY 215
|||||
DB 378 GSSVLVLY 385

RESULT 15

T50309

hypothetical WD-repeat protein [imported] - fission yeast (Schizosaccharomyces pombe)

C/Species: Schizosaccharomyces pombe

C/Date: 09-Jun-2000 #sequence_revision 09-Jun-2000 #text_change 09-Jul-2004

C/Accession: T50309

R/Brown, S.; Harris, D.; McDougall, R.C.; Rejandream, M.A.; Barrell, B.G. submitted to the EMBL Data Library, November 1999

A/Reference number: Z25060

A/Accession: T50309

A/Status: preliminary; translated from GB/EMBL/DBJ

A/Molecule type: DNA

A/Residues: 1-384 <BRO>

A/Cross-references: UNIPROT:Q9USZ0; EMBL:AL133303; PIDN:CA862092.1; GSPDB:GN00067; SPDB:G

A/Experimental source: strain 972h(-); cosmid c1306

C/Genetics:

A/Gene: SPBC4.08; SPDB:SPBC1306.02

A/Map position: 2

A/Introns: 349/2; 565/3

Query Match 2.4%; Score 8; DB 2; Length 984;
Best Local Similarity 100.0%; Pred. No. 19;
Matches 8; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 280 SSANPIIY 287
|||||
DB 290 SSANPIIY 297

Search completed: October 27, 2004, 08:53:08
Job time : 41 secs

GenCore version 5.1.6
Copyright (c) 1993 - 2004 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: October 27, 2004, 08:35:09 ; Search time 196 Seconds

(without alignments)
989.292 Million cell updates/sec

Title: US-09-867-570-2

Sequence: 1 MSKSKSWVRLGFLMSDSTI.....EGSGMLPQETLBSGRLAQ 337

Scoring table: OLIGO
Gapop 60.0 , Gapext 60.0

Searched: 1825181 seqs, 575374646 residues

Word size : 0

Total number of hits satisfying chosen parameters: 1825181

Minimum DB seq length: 0

Maximum DB seq length: 200000000

Post-processing: Listing first 45 summaries

Database :

1: uniprot_02:*
2: uniprot_trembl:*

Pred. No. is the number of results predicted by chance to have a
score greater than or equal to the score of the result being printed,
and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	322	95.5	322	MRG3_HUMAN	Q961B0 homo sapien
2	224	66.5	322	AAH67292	AAH67292 homo sapi
3	86	25.5	322	SNS2_HUMAN	Q8TDE0 homo sapien
4	50	14.8	322	MRG1_HUMAN	Q961B2 homo sapien
5	44	13.1	322	SNS3_HUMAN	Q8TDE9 homo sapien
6	41	12.2	322	MRG4_HUMAN	Q961A9 homo sapien
7	41	12.2	322	SNS5_HUMAN	Q8TDE7 homo sapien
8	20	5.9	330	MRG2_HUMAN	Q961B1 homo sapien
9	20	5.9	330	AAH63450	AAH63450 homo sapi
10	15	4.5	302	MG43_MOUSE	Q91WV3 mus musculu
11	15	4.5	304	MG41_MOUSE	Q91WV5 mus musculu
12	15	4.5	304	MG42_MOUSE	Q91WV6 mus musculu
13	15	4.5	305	MG47_MOUSE	Q91ZC5 mus musculu
14	15	4.5	305	MG48_MOUSE	Q91ZC4 mus musculu
15	15	4.5	313	MG49_MOUSE	Q91W2 mus musculu
16	15	4.5	313	MG47_MOUSE	Q91W2 mus musculu
17	15	4.5	313	MG47_MOUSE	Q91W2 mus musculu
18	14	4.2	378	MRG_HUMAN	Q91YB7 ratu
19	13	3.9	301	AAH69345	AAH69345 homo sapi
20	13	3.9	301	MG46_MOUSE	Q91ZC6 mus musculu
21	13	3.9	305	MG42_MOUSE	Q91W2 mus musculu
22	11	3.3	331	AAH64040	AAH64040 mus muscu
23	11	3.3	331	AAH64040	AAH64040 mus muscu
24	11	3.3	331	AAH64040	AAH64040 mus muscu
25	11	3.3	331	AAH64040	AAH64040 mus muscu
26	11	3.3	331	AAH64040	AAH64040 mus muscu
27	10	3.0	312	Q91ZC1	Q91ZC1 mus musculu
28	10	3.0	312	Q91ZC1	Q91ZC1 mus musculu
29	10	3.0	312	Q91ZC1	Q91ZC1 mus musculu
30	10	3.0	312	Q91ZC1	Q91ZC1 mus musculu
31	9	2.7	321	Q7TN38	Q7TN38 ratu

32	9	2.7	353	2	Q7TN44	Q7TN44 ratu
33	9	2.7	399	2	Q75XU5	Q75XU5 gallu
34	9	2.7	399	2	BAC67782	BAC67782 gallu
35	9	2.7	430	1	NFPI_HUMAN	Q99ZG6 homo sapien
36	8	2.4	111	1	THN3_VISAL	P01538 viscu
37	8	2.4	116	2	Q9HBE6	Q9HBE6 homo sapien
38	8	2.4	116	2	Q96H14	Q96H14 homo sapien
39	8	2.4	141	2	Q9JY80	Q9JY80 neisseria m
40	8	2.4	141	2	Q9JY98	Q9JY98 neisseria m
41	8	2.4	156	2	Q711N2	Q711N2 mus musculu
42	8	2.4	156	2	CAC68130	CAC68130 mus muscu
43	8	2.4	164	2	Q82858	Q82858 jemrana d1
44	8	2.4	215	1	SOMA_RANCA	P10813 rana catesb
45	8	2.4	215	2	Q7ZU47	Q7ZU47 rana catesb

ALIGNMENTS

RESULT 1
MRG3_HUMAN STANDARD; PRT; 322 AA.
ID MRG3_HUMAN
AC Q961B0; Q8TDE1;
DT 01-OCT-2004 (Rel. 45, Created)
DT 01-OCT-2004 (Rel. 45, Last sequence update)
DT 01-OCT-2004 (Rel. 45, Last annotation update)
DE Mas-related G-protein coupled receptor member X3 (Sensory neuron-specific G-protein coupled receptor 1).
GN Name=MRG3; Synonyms=SNR1;
OS Homo sapiens (human).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.
NCBI_TaxID=9606;
RX MDL:LINE=21853733; PubMed=11850634; DOI=10.1038/nr15;
RA Lembo P.M.C., Grazzini B., Groblewski T., O'Donnell D., Roy M.-O., Zhang J., Hoffert C., Cao J., Schmidt R., Pelletier M., Labarre M., Gosselin M., Fortin Y., Barville D., Shen S., Stroem P., Payza K., Dray A., Walker P., Ahmad S.;
RT "Proenkephalin A gene products activate a new family of sensory neuron-specific GPCRs.";
RL Nat. Neurosci. 5:201-209(2002).
RN [3]
RP SEQUENCE FROM N.A., VARIANT ASN-169, AND TISSUE SPECIFICITY.
RC TISSUE=Testis;
RX MDL:LINE=22388257; PubMed=12477932; DOI=10.1073/pnas.242603899;
RA Krausberg R.L., Feingold E.A., Grouse L.H., Derge J.G.,
RA Krausberg R.L., Collins P.S., Wagner L., Shennan C.M., Schuler G.D.,
RA Altschul S.F., Zeeberg B., Buettow K.H., Schaefer C.F., Bhat N.K.,
RA Hopkin R.F., Jordan H., Moore T., Wax S.T., Wang J., Hsieh P.,
RA Diatchenko L., Marusik A., Farmer A.A., Rubin G.M., Hong L.,
RA Stampstein M., Soares M.B., Bonaldi M.F., Casavant T.L., Scheetz T.B.,
RA Brownstein M.J., Uedlin T.B., Tohiyuki S., Carninci P., Prange C.,
RA Raba S.S., Liguori N.A., Peters G.J., Abramson R.D., Mulhally S.J.,
RA Bosak S.A., McEwan P.J., McKernan K.J., Malek J.A., Gunaratne P.H.,
RA Richards S., Worley K.C., Hale S., Garcia A.M., Gay L.D., Hulyk S.W.,
RA Vallian D.K., Muzny D.M., Sodergren E.J., Lu X., Gibbs R.A.,
RA Fahy J., Helton E., Kettelman M., Madan A., Rodriguez S., Sanchez A.,
RA Whiting M., Madan A., Young A.C., Shevchenko Y., Bouffard G.G.,
RA Blakesley R.W., Touchman J.W., Green E.D., Dickson R.M.,
RA Rodriguez A.C., Grimwood J., Schmutz J., Myers R.M.,
RA Butterfield Y.S.N., Krzywinski M.I., Skalska U., Smalins D.B.,
RA Schnerch A., Schein J.E., Jones S.J.M., Maria M.A.,
RT "Generation and initial analysis of more than 15,000 full-length human
and mouse cDNA sequences.";

RL Proc. Natl. Acad. Sci. U.S.A. 99:16899-16903(2002).
 CC -1- FUNCTION: Orphan receptor. Probably involved in the function of nociceptive neurons. May regulate nociceptor function and/or development, including the sensation or modulation of pain.
 CC Potentially activated by enkephalins (By similarity).
 CC -1- SUBCELLULAR LOCATION: Integral membrane protein.
 CC -1- TISSUE SPECIFICITY: Uniquely localized in a subset of small dorsal root and trigeminal sensory neurons.
 CC -1- SIMILARITY: Belongs to family 1 of G-protein coupled receptors. Mas subfamily.
 CC -----
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration between the Swiss Institute of Bioinformatics and the EMBL outstation - the European Bioinformatics Institute. There are no restrictions on its use by non-profit institutions as long as its content is in no way modified and this statement is not removed. Usage by and for commercial entities requires a license agreement (See <http://www.isb-sib.ch/announce/> or send an email to license@isb-sib.ch).
 CC -----
 CC EMBL; AY042215; AAK91806.1; -
 CC EMBL; AF474987; AAL6878.2; -
 CC EMBL; BC067292; AAK67292.1; -
 CC MIM; 607229; -
 CC InterPro; IPR000276; GPCR_Rhodopsn.
 CC Pfam; PF00001; 7tm_1; 1.
 CC PRINTS; PR00237; GPCR_Rhodopsn.
 CC PROSITE; PS00237; G-PROTEIN_RECEP_F1_1; 1.
 CC PROSITE; PS50262; G-PROTEIN_RECEP_F1_2; 1.
 CC G-protein coupled receptor; Polymorphism; Transmembrane.
 CC DOMAIN 1 31 Extracellular (Potential).
 CC TRANSSEM 32 52 1 (Potential).
 CC TRANSSEM 53 60 Cytoplasmic (Potential).
 CC TRANSSEM 61 81 2 (Potential).
 CC TRANSSEM 82 96 Extracellular (Potential).
 CC TRANSSEM 97 117 3 (Potential).
 CC TRANSSEM 118 140 Cytoplasmic (Potential).
 CC TRANSSEM 141 161 4 (Potential).
 CC TRANSSEM 162 177 Extracellular (Potential).
 CC TRANSSEM 178 198 5 (Potential).
 CC TRANSSEM 199 213 Cytoplasmic (Potential).
 CC TRANSSEM 214 234 6 (Potential).
 CC TRANSSEM 235 254 Extracellular (Potential).
 CC TRANSSEM 255 275 7 (Potential).
 CC TRANSSEM 276 322 Cytoplasmic (Potential).
 CC TRANSSEM 169 169 D -> N (in dbSNP:4274188).
 CC VARIANT /FTID=VAR 019434.
 CC CONFLICT 3 3 S -> P (in Ref. 2).
 CC CONFLICT 82 82 C -> R (in Ref. 3).
 CC CONFLICT 307 307 W -> Q (in Ref. 3).
 CC CONFLICT 319 319 R -> K (in Ref. 2).
 CC SEQUENCE 322 AA; 36484 MW; 2531BBF0CB4EB74 CRC64;
 SQ
 Query Match 95.5%; Score 322; DB 1; Length 322;
 Best Local Similarity 100.0%; Pred. No. 3e-111;
 Matches 322; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

DB 241 RHLDKVLVFCVHLVSIPLSALNSSANPIYFVGSFRORONKLVLRALQDTPPE 300
 QY 316 VDEGGWLPQETLESGSRLEQ 337
 DB 301 VDEGGWLPQETLESGSRLEQ 322
 RESULT 2
 AAH67292
 ID AAH67292 PRELIMINARY; PRT; 322 AA.
 AC AAH67292;
 DT 25-MAR-2004 (TRENBLREL. 27, Created)
 DT 25-MAR-2004 (TRENBLREL. 27, Last sequence update)
 DT 25-MAR-2004 (TRENBLREL. 27, Last annotation update)
 DE G-protein-coupled receptor MRGX3.
 GN MRGX3.
 OS Homo sapiens (Human).
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 OC Mammalia; Eutheria; Primates; Catarrhini; Homiidae; Homo.
 OX NCBI_Taxid=9606;
 RN [1]
 RC SEQUENCE FROM N.A.
 RP TISSUE=Testis;
 RX MEDLINE=22388257; PubMed=12477932;
 RA Strassberg R.L., Feingold E.A., Grouse L.H., Derge J.G.,
 RA Klausner R.D., Collins F.S., Wagner L., Shennan C.M., Schuler G.D.,
 RA Altschul S.F., Zeeberg B., Buetow K.H., Schaefer C.F., Hlat N.K.,
 RA Hopkins R.F., Jordan H., Moore T., Max S.I., Wang J., Hlat F.,
 RA Diatchenko L., Marinina K., Farmer A.A., Rubin G.M., Hong L.,
 RA Stapleton M., Soares M.B., Bonaldo M.F., Casavant T.L., Scheetz T.E.,
 RA Brownstein M.J., Usdin T.B., Toshiyuki S., Carninci P., Prange C.,
 RA Raha S.S., Loggiano N.A., Peters G.J., Abramson R.D., Mullany S.J.,
 RA Bosak S.A., McKean P.J., McKernan K.J., Malek J.A., Gunaratne P.H.,
 RA Richards S., Morley K.C., Hale S., Garcia A.M., Gay L.J., Hulyk S.W.,
 RA Villalón D.K., Muzny D.M., Sodergren E.J., Lu X., Gibbs R.A.,
 RA Fahney J., Helton K., Kettelman M., Madan A., Rodriguez S., Sanchez A.,
 RA Whitting M., Madan A., Young A.C., Shevchenko Y., Boultard G.G.,
 RA Blakesley R.W., Touchman J.W., Green E.D., Dickson M.C.,
 RA Rodriguez A.C., Grimwood J., Schmutz J., Myers R.M., Butterfield Y.S.,
 RA Krzyzanski M.I., Skalske U., Smallos D.B., Schnerch A., Schein J.B.,
 RA Jones S.J., Merra M.A.;
 RA "Generation and initial analysis of more than 15,000 full-length human
 RT and mouse cDNA sequences";
 RL Proc. Natl. Acad. Sci. U.S.A. 99:16899-16903(2002).
 RN [2]
 RP SEQUENCE FROM N.A.
 RC TISSUE=Testis;
 RA Strassberg R.;
 RL Submitted (MAR-2004) to the EMBL/GenBank/DBJ databases.
 DR EMBL; BC067292; AAK67292.1; -
 KW Receptor.
 SQ SEQUENCE 322 AA; 36479 MW; B91DC082B6D95DA8 CRC64;
 Query Match 66.5%; Score 224; DB 2; Length 322;
 Best Local Similarity 100.0%; Pred. No. 9e-214;
 Matches 224; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

```

RESULT 3
SNS2_HUMAN          STANDARD;          PRT;          322 AA.
ID   SNS2_HUMAN      O8TDE0;
DT   01-OCT-2004 (Rel. 45, Created)
DT   01-OCT-2004 (Rel. 45, Last sequence update)
DT   01-OCT-2004 (Rel. 45, Last annotation update)
DE   Sensory neuron-specific G-protein coupled receptor 2.
GN   Name=SNSR2;
OS   Homo sapiens (Human).
OC   Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi;
OC   Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.
OC   NCBI_TaxId=9606;
RN   [1]
RP   MEDLINE=21853733; PubMed=11850634; DOI=10.1038/nr815;
RA   Lembo P.M.C., Grazzini E., Groblewski T., O'Donnell D., Roy M.-O.,
RA   Zhang J., Hoffert C., Cao J., Schmidt R., Pelletier M., Labarre M.,
RA   Gosselin M., Fortin Y., Banville D., Shen S., Stroem P., Payza K.,
RA   Dray A., Walker P., Ahmad S.;
RT   "Proenkephalin A gene products activate a new family of sensory
RT   neuron-specific GPCRs."
RL   Nat. Neurosci. 5:201-209(2002).
CC   -1- FUNCTION: Orphan receptor. Probably involved in the function of
CC   nociceptive neurons. May regulate nociceptor function and/or
CC   development, including the sensation or modulation of pain.
CC   Potentially activated by enkephalins (By similarity).
CC   -1- SUBCELLULAR LOCATION: Integral membrane protein.
CC   -1- TISSUE SPECIFICITY: Uniquely localized in a subset of small dorsal
CC   root and trigeminal sensory neurons.
CC   -1- SIMILARITY: Belongs to family 1 of G-protein coupled receptors.
CC   Was subfamily.
-----
CC   This SWISS-PROT entry is copyright. It is produced through a collaboration
CC   between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC   the European Bioinformatics Institute. There are no restrictions on its
CC   use by non-profit institutions as long as its content is in no way
CC   modified and this statement is not removed. Usage by and for commercial
CC   entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC   or send an email to license@isb-sib.ch).
-----
EMBL: AF474988; AAL86879.2; -.
DR   InterPro: IPR000276; GPCR_Rhodopsin.
DR   Pfam: PF00001; 7cm_1, 1.
DR   PRINTS: PR00237; GPCRHHODOPS.
DR   PROSITE: PS00237; G_PROTEIN_RECER_F1_1; 1.
DR   PROSITE: PS50262; G_PROTEIN_RECER_F2_1; 1.
KW   G-protein coupled receptor; Transmembrane.
FT   DOMAIN          1      322
FT   TRANSMEM       33      53
FT   DOMAIN          54      60
FT   TRANSMEM       61      81
FT   DOMAIN          82      96
FT   TRANSMEM       97      117
FT   DOMAIN          118     140
FT   TRANSMEM       141     161
FT   DOMAIN          162     177
FT   TRANSMEM       178     198
FT   TRANSMEM       199     213
FT   TRANSMEM       214     234
FT   DOMAIN          235     254
FT   TRANSMEM       255     275
FT   DOMAIN          276     322
FT   CARBOHYD       89
FT   SEQUENCE       322 AA; 36594 MW; D8C24308B34611B CRC64;
SQ
Query Match          25.5%; Score 86; DB 1; Length 322;
Best Local Similarity 100.0%; Pred. No. 1.7e-76;
Matches 86; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
204 VVLGSSLVVLVRLTLCGRKMPILRLVVTLLVTLVFLGCLPGIGIOWALFSLRHDMKV 263

```

```

Db          189
VVLGSSLVVLVRLTLCGRKMPILRLVVTLLVTLVFLGCLPGIGIOWALFSLRHDMKV 248
Qy          264 LFCVHLVSLFLSALNSSANPITYFF 289
Db          249 LFCVHLVSLFLSALNSSANPITYFF 274

RESULT 4
MRG1_HUMAN          STANDARD;          PRT;          322 AA.
ID   MRG1_HUMAN      Q96LB2; Q8TDB;
DT   01-OCT-2004 (Rel. 45, Created)
DT   01-OCT-2004 (Rel. 45, Last sequence update)
DT   01-OCT-2004 (Rel. 45, Last annotation update)
DE   Mas-related G-protein coupled receptor member X1 (Sensory neuron-
DE   specific G-protein coupled receptor 4).
GN   Name=MRGX1; Synonym=SNSR4;
OS   Homo sapiens (Human).
OC   Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC   Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.
OC   NCBI_TaxId=9606;
RN   [1]
RP   MEDLINE=21435808; PubMed=11551509;
RA   Dong X., Han S.-K., Zylka M.J., Simon M.I., Anderson D.J.;
RA   "A diverse family of GPCRs expressed in specific subsets of
RA   nociceptive sensory neurons."
RL   Cell 106:619-632(2001).
RN   [2]
RP   MEDLINE=21853733; PubMed=11850634; DOI=10.1038/nr815;
RA   Lembo P.M.C., Grazzini E., Groblewski T., O'Donnell D., Roy M.-O.,
RA   Zhang J., Hoffert C., Cao J., Schmidt R., Pelletier M., Labarre M.,
RA   Gosselin M., Fortin Y., Banville D., Shen S., Stroem P., Payza K.,
RA   Dray A., Walker P., Ahmad S.;
RT   "Proenkephalin A gene products activate a new family of sensory
RT   neuron-specific GPCRs."
RL   Nat. Neurosci. 5:201-209(2002).
RN   [3]
RP   MEDLINE=22040266; PubMed=12044878;
RA   Takeeda S., Kadowaki S., Haga T., Takeasu H., Mitaku S.;
RT   "Identification of G protein-coupled receptor genes from the human
RT   genome sequence."
RL   FEMS Lett. 520:97-101(2002).
RN   [4]
RP   SEQUENCE FROM N.A.
RA   Suwa M., Sato T., Okouchi I., Arita M., Futami K., Matsumoto S.,
RA   Tautani S., Aburatani H., Asai K., Akiyama Y.;
RT   "genome-wide discovery and analysis of human seven transmembrane helix
RT   receptor genes."
RL   Submitted (JUL-2001) to the EMBL/GenBank/DBJ databases.
CC   -1- FUNCTION: Orphan receptor. Probably involved in the function of
CC   nociceptive neurons. May regulate nociceptor function and/or
CC   development, including the sensation or modulation of pain.
CC   Potentially activated by enkephalins including BAM2 (bovine adrenal
CC   medulla peptide 22) and BAM (8-22). BAM22 is the most potent
CC   compound and evoked a large and dose-dependent release of
CC   intracellular calcium in stably transfected cells. G(alpha)q
CC   proteins are involved in the calcium-signaling pathway.
CC   -1- SUBCELLULAR LOCATION: Integral membrane protein.
CC   -1- TISSUE SPECIFICITY: Uniquely localized in a subset of small dorsal
CC   root and trigeminal sensory neurons.
CC   -1- SIMILARITY: Belongs to family 1 of G-protein coupled receptors.
CC   Was subfamily.
-----
CC   This SWISS-PROT entry is copyright. It is produced through a collaboration
CC   between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC   the European Bioinformatics Institute. There are no restrictions on its
CC   use by non-profit institutions as long as its content is in no way
CC   modified and this statement is not removed. Usage by and for commercial
CC   entities requires a license agreement (See http://www.isb-sib.ch/announce/

```


RA Lembo P.M.C., Grazzini E., Groblewski T., O'Donnell D., Roy M.-O.,
 RA Zhang J., Hoffert C., Cao J., Schmidt R., Pelletier M., Labarre M.,
 RA Gosselin M., Fortin Y., Banville D., Shen S., Stroem P., Payza K.,
 RA Dray A., Walker P., Ahmad S.,
 RT "Proenkephalin A gene products activate a new family of sensory
 RT neuron-specific GPCRs.",
 RL Nat. Neurosci. 5:201-209(2002).
 CC -1- FUNCTION: Orphan receptor. Probably involved in the function of
 CC nociceptive neurons. May regulate nociceptor function and/or
 CC development, including the sensation or modulation of pain.
 CC Potentially activated by enkephalins (By similarity).
 CC -1- SUBCELLULAR LOCATION: Integral membrane protein.
 CC -1- TISSUE SPECIFICITY: Uniquely localized in a subset of small dorsal
 CC root and trigeminal sensory neurons.
 CC -1- SIMILARITY: Belongs to family 1 of G-protein coupled receptors.
 CC Was subfamily.
 CC -----
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
 CC the European Bioinformatics Institute. There are no restrictions on its
 CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
 CC or send an email to license@isb-sib.ch).
 CC -----
 DR EMBL: AF42216; AAK91807.1; -
 DR EMBL: AF474992; AAL6883.1; -
 DR MTM: 607230; -
 DR InterPro: IPR00276; GPCR_Rhodopsn.
 DR Pfam: PF00001; 7tm.1; 1.
 DR PRINTS: PR00237; GPCR_RHODOPSN.
 DR PROSITE: PS00237; G_PROTEIN_REC_P1_1; 1.
 DR PROSITE: PS50262; G_PROTEIN_REC_P2_1; 1.
 KW G-protein coupled receptor; Glycoprotein; Polymorphism; Transmembrane.
 FT DOMAIN 1 31
 FT TRANSSEM 32 52
 FT TRANSSEM 53 60
 FT TRANSSEM 61 81
 FT DOMAIN 82 96
 FT TRANSSEM 97 117
 FT DOMAIN 118 137
 FT TRANSSEM 138 158
 FT DOMAIN 159 177
 FT TRANSSEM 178 198
 FT DOMAIN 199 218
 FT TRANSSEM 219 239
 FT DOMAIN 240 254
 FT TRANSSEM 255 275
 FT DOMAIN 276 322
 FT CARBOHYD 25 25
 FT CARBOHYD 89 89
 FT VARIANT 8 8
 FT VARIANT 25 25
 FT VARIANT 54 54
 FT VARIANT 83 83
 FT CONFLICT 182 182
 FT CONFLICT 319 319
 FT SEQUENCE 322 AA; 36434 MW; 7CA676F8BD30A1 CRC64;
 Query Match 12.2%; Score 41; DB 1; Length 322;
 Best Local Similarity 100.0%; Pred. No. 9.6e-32;
 Matches 41; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
 QY 209 SSVLVLRILGSRKMPRLTYTLLTVLVFLGSPFGI 249
 DB 194 SSVLVLRILGSRKMPRLTYTLLTVLVFLGSPFGI 234

SNSS HUMAN
 ID SNSS HUMAN STANDARD; PRT; 322 AA.
 DT 01-OCT-2004 (Rel. 45, Created)
 DT 01-OCT-2004 (Rel. 45, Last sequence update)
 DT 01-OCT-2004 (Rel. 45, Last annotation update)
 DE Sensory neuron-specific G-protein coupled receptor 5.
 GN Names: SNRS;
 OS Homo sapiens (Human).
 OC Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi;
 OC Mammalia; Eutheria; Primates; Catarrhini; Homiidae; Homo.
 OC NCBI_TaxID=9606;
 RN [1]
 RP SEQUENCE FROM N.A. AND TISSUE SPECIFICITY.
 RX MEDLINE=11853733; PubMed=11850634; DOI=10.1038/nn815;
 RA Lembo P.M.C., Grazzini E., Groblewski T., O'Donnell D., Roy M.-O.,
 RA Zhang J., Hoffert C., Cao J., Schmidt R., Pelletier M., Labarre M.,
 RA Gosselin M., Fortin Y., Banville D., Shen S., Stroem P., Payza K.,
 RA Dray A., Walker P., Ahmad S.,
 RT "Proenkephalin A gene products activate a new family of sensory
 RT neuron-specific GPCRs.",
 RL Nat. Neurosci. 5:201-209(2002).
 CC -1- FUNCTION: Orphan receptor. Probably involved in the function of
 CC nociceptive neurons. May regulate nociceptor function and/or
 CC development, including the sensation or modulation of pain.
 CC Potentially activated by enkephalins (By similarity).
 CC -1- SUBCELLULAR LOCATION: Integral membrane protein.
 CC -1- TISSUE SPECIFICITY: Uniquely localized in a subset of small dorsal
 CC root and trigeminal sensory neurons.
 CC -1- SIMILARITY: Belongs to family 1 of G-protein coupled receptors.
 CC Was subfamily.
 CC -----
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
 CC the European Bioinformatics Institute. There are no restrictions on its
 CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
 CC or send an email to license@isb-sib.ch).
 CC -----
 DR EMBL: AF474991; AAL6882.1; -
 DR InterPro: IPR00276; GPCR_Rhodopsn.
 DR Pfam: PF00001; 7tm.1; 1.
 DR PRINTS: PR00237; GPCR_RHODOPSN.
 DR PROSITE: PS00237; G_PROTEIN_REC_P1_1; 1.
 DR PROSITE: PS50262; G_PROTEIN_REC_P2_1; 1.
 KW G-protein coupled receptor; Transmembrane.
 FT DOMAIN 1 31
 FT TRANSSEM 32 52
 FT TRANSSEM 53 60
 FT TRANSSEM 61 81
 FT DOMAIN 82 96
 FT TRANSSEM 97 117
 FT DOMAIN 118 137
 FT TRANSSEM 138 158
 FT DOMAIN 159 177
 FT TRANSSEM 178 198
 FT DOMAIN 199 218
 FT TRANSSEM 219 239
 FT DOMAIN 240 254
 FT TRANSSEM 255 275
 FT DOMAIN 276 322
 FT CARBOHYD 89 89
 FT SEQUENCE 322 AA; 36423 MW; 3D6FBA5DDFD90 CRC64;
 Query Match 12.2%; Score 41; DB 1; Length 322;
 Best Local Similarity 100.0%; Pred. No. 9.6e-32;
 Matches 41; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
 QY 209 SSVLVLRILGSRKMPRLTYTLLTVLVFLGSPFGI 249
 DB 194 SSVLVLRILGSRKMPRLTYTLLTVLVFLGSPFGI 234

RA MEDLINE=22388257; PubMed=12477932;
RA Strausberg R.L., Feingold B.A., Grouse L.H., Derge J.G.,
RA Klausner R.D., Collins F.S., Wagner L., Shemen C.M., Schlier G.D.,
RA Altschul S.F., Zeeberg B., Buetow K.H., Schaefer C.F., Bhat N.K.,
RA Hopkins R.F., Jordan H., Moore T., Max S.I., Wang J., Hsieh F.,
RA Diatchenko L., Marusik A.K., Farmer A.A., Rubin G.M., Hong L.,
RA Scapleorn M.J., Soares M.B., Bonaldo M.F., Casavani P., Prange C.,
RA Brownstein M.J., Uadit T.B., Tohyuki S., Carninci P., Scheetz T.E.,
RA Raha S.S., Loggellano N.A., Peters G.J., Abramson R.D., Mullany S.J.,
RA Bosak S.A., McSwan P.J., McKernan K.J., Malek J.A., Gunaratne P.H.,
RA Richards S., Worley K.C., Hale S., Garcia A.M., Gay L.J., Huiyk S.W.,
RA Villalon D.K., Muzny D.M., Sodergren E.J., Lu X., Gibbs R.A.,
RA Fahy J., Helton E., Kettelman M., Madan A., Rodriguez S., Sanchez A.,
RA Whiting M., Madan A., Young A.C., Shevchenko Y., Bouffard G.G.,
RA Blakesley R.W., Touchman J.W., Green E.D., Dickson M.C.,
RA Rodriguez A.C., Grimwood J., Schmutz J., Myers R.M., Butterfield Y.S.,
RA Krzywiński M.I., Skalka U., Smallin D.E., Schnerch A., Schein J.E.,
RA Jones S.J., Marra M.A.;
RT "Generation and initial analysis of more than 15,000 full-length human
RT and mouse cDNA sequences.";
RT Proc. Natl. Acad. Sci. U.S.A. 99:16899-16903(2002).
RN (2)
RP SEQUENCE FROM N.A.
RC TISSUE=Brain;
RA Strausberg R.;
RL Submitted (DEC-2003) to the EMBL/GenBank/DBJ databases.
DR EMBL, BC063450; AA063450.1; -.
KW Receptor.
SQ SEQUENCE 330 AA; 37099 MW; 0B328FD78B1DF6BE CRC64;
Query Match 5.9%; Score 20; DB 2; Length 330;
Best Local Similarity 100.0%; Pred. No. 7.6e-11;
Matches 20; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
Qy 232 TILLTVLPFLCGLPFGIQW 251
Db 224 TILLTVLPFLCGLPFGIQW 243

RESULT 10
MG3 MOUSE STANDARD; PRT; 302 AA.
AC 091MW3;
DT 01-OCT-2004 (Rel. 45, Last sequence update)
DT 01-OCT-2004 (Rel. 45, Last annotation update)
DE Mas-related G-protein coupled receptor member A3.
GN Name=Mrgpr3; Synonyms=Mrga3;
OS Mus musculus (Mouse).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
OX NCBI_TaxID=10090;
RN (1)
RP SEQUENCE FROM N.A., AND TISSUE SPECIFICITY.
RC STRAIN=C57BL/6; TISSUE=Dorsal root ganglion;
RX MEDLINE=21435808; PubMed=1151509;
RA Dong X., Han S.-K., Zylka M.J., Simon M.I., Anderson D.J.;
RT "A diverse family of GPCRs expressed in specific subsets of
RT nociceptive sensory neurons.";
RL Cell 106:619-632(2001).
-1- FUNCTION: Orphan receptor. May be a receptor for Rfamide-family
neuropeptides such as NPFF and NPAF, which are analgesic in vivo.
May regulate nociceptor function and/or development, including the
sensation or modulation of pain (By similarity).
-1- SUBCELLULAR LOCATION: Integral membrane protein.
-1- TISSUE SPECIFICITY: Expressed in a subset of sensory neurons that
includes nociceptors. Expressed in the subclasses of nonpeptidergic
sensory neurons that are IB4(+) and VR1(-).
-1- SIMILARITY: Belongs to family 1 of G-protein coupled receptors.
between the Swiss Institute of Bioinformatics and the EMBL outstation -

CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See <http://www.ebi.ac.uk/announce/>
CC or send an email to license@ebi.ac.uk).
DR EMBL, AY042193; AK91789.1; -.
DR MGD; AY042193; AK91789.1; -.
DR InterPro: IPR000276; GPCR_Rhodopsin.
DR Pfam: PF00001; 7tm1.1; 1.
DR PRINTS; PR00237; GPCRHOOPS.
DR PROSITE; PS00237; G_PROTEIN_RECEP_F1_1; 1.
DR PROSITE; PS00262; G_PROTEIN_RECEP_F1_2; 1.
KM G-protein coupled receptor; Glycoprotein; Transmembrane.
FT DOMAIN 1 17
FT TRANSMEM 18 36
FT DOMAIN 39 46
FT TRANSMEM 47 67
FT DOMAIN 68 81
FT TRANSMEM 82 102
FT DOMAIN 103 129
FT TRANSMEM 130 150
FT DOMAIN 151 167
FT TRANSMEM 168 188
FT DOMAIN 189 211
FT TRANSMEM 212 232
FT DOMAIN 233 242
FT TRANSMEM 243 263
FT DOMAIN 264 302
FT CARBOHYD 2 2
FT CARBOHYD 68 68
FT CARBOHYD 159 159
SQ SEQUENCE 302 AA; 34483 MW; 8E4CC023AC2B5E5 CRC64;
Query Match 4.5%; Score 15; DB 1; Length 302;
Best Local Similarity 100.0%; Pred. No. 6.6e-06;
Matches 15; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
Qy 124 GLSMIASTERCLS 138
Db 98 GLSMIASTERCLS 112

RESULT 11
MG1 MOUSE STANDARD; PRT; 304 AA.
AC 091MW5;
DT 01-OCT-2004 (Rel. 45, Last sequence update)
DT 01-OCT-2004 (Rel. 45, Last annotation update)
DE Mas-related G-protein coupled receptor member A1 (Rf-amide G protein-
DE coupled receptor).
GN Name=Mrgpr1; Synonyms=Mrga1;
OS Mus musculus (Mouse).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
OX NCBI_TaxID=10090;
RN (1)
RP SEQUENCE FROM N.A., SUBCELLULAR LOCATION, AND TISSUE SPECIFICITY.
RC STRAIN=C57BL/6; TISSUE=Dorsal root ganglion;
RX MEDLINE=21435808; PubMed=1151509;
RA Dong X., Han S.-K., Zylka M.J., Simon M.I., Anderson D.J.;
RT "A diverse family of GPCRs expressed in specific subsets of
RT nociceptive sensory neurons.";
RL Cell 106:619-632(2001).
-1- FUNCTION: Orphan receptor. May be a receptor for Rfamide-family
neuropeptides such as NPFF and NPAF, which are analgesic in vivo.
May regulate nociceptor function and/or development, including the
sensation or modulation of pain.
-1- SUBCELLULAR LOCATION: Integral membrane protein.
-1- TISSUE SPECIFICITY: Expressed in a subset of sensory neurons that
includes nociceptors. Expressed in the subclasses of nonpeptidergic
sensory neurons that are IB4(+) and VR1(-).
between the Swiss Institute of Bioinformatics and the EMBL outstation -

CC -1- SIMILARITY: Belongs to family 1 of G-protein coupled receptors.
CC Was subfamily.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
CC or send an email to license@isb-sib.ch).
CC -----
DR EMBL, AY042191; AAK91787.1; -.
DR WCD; MG1:3033095; Mrgpra1.
DR InterPro; IPR000276; GPCR_Rhodpsn.
DR Pfam; PF00001; 7tm 1; 1.
DR PRINTS; PR00237; GPCRHOOPS.
DR PROSITE; PS00237; G_PROTEIN_RECEP_F1_1; 1.
DR PROSITE; PS50262; G_PROTEIN_RECEP_F1_2; 1.
KW G-protein coupled receptor; Glycoprotein; Transmembrane.
FT DOMAIN 1 17 Extracellular (Potential).
FT TRANSMEM 18 38 Cytoplasmic (Potential).
FT DOMAIN 39 53 Cytoplasmic (Potential).
FT TRANSMEM 54 74 Extracellular (Potential).
FT DOMAIN 75 75 Extracellular (Potential).
FT TRANSMEM 76 96 Cytoplasmic (Potential).
FT DOMAIN 97 131 Cytoplasmic (Potential).
FT TRANSMEM 132 152 Extracellular (Potential).
FT DOMAIN 153 166 Extracellular (Potential).
FT TRANSMEM 167 187 Cytoplasmic (Potential).
FT DOMAIN 188 206 Cytoplasmic (Potential).
FT TRANSMEM 207 227 Extracellular (Potential).
FT DOMAIN 228 243 Extracellular (Potential).
FT TRANSMEM 244 264 Cytoplasmic (Potential).
FT DOMAIN 265 304 N-linked (GlcNAc) (Potential).
FT CARBOHYD 10 10 N-linked (GlcNAc) (Potential).
SQ SEQUENCE 304 AA; 34381 MW; C56CBF879067A52B CRC64;

Query Match 4.5%; Score 15; DB 1; Length 304;
Best Local Similarity 100.0%; Pred. No. 6.7e-06;
Matches 15; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 124 GLSMLSAISTERCLS 138
Db 97 GLSMLSAISTERCLS 111

RESULT 12
MRCG_RAT STANDARD; PRT; 304 AA.
ID MRCG_RAT STANDARD; PRT; 304 AA.
AC Q7TN49;
DT 01-OCT-2004 (Rel. 45, Created)
DT 01-OCT-2004 (Rel. 45, Last sequence update)
DT 01-OCT-2004 (Rel. 45, Last annotation update)
DE Mas-related G-protein coupled receptor member A.
GN Name=Mrgpra; Synonyms=Mrga;
OS Rattus norvegicus (Rat).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Rattus.
OX NCBI_TaxID=10116;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=Sprague-Dawley;
RX MEDLINE=22810130; PubMed=12909716; DOI=10.1073/pnas.1732949100;
RA Zylka M.J., Dong X., Southwell A.L., Anderson D.J.;
RT "Atypical expansion in mice of the sensory neuron-specific Mrg G
RT protein-coupled receptor family."
RL Proc. Natl. Acad. Sci. U.S.A. 100:10043-10048(2003).
CC -1- FUNCTION: Orphan receptor. May regulate nociceptor function and/or
CC development, including the sensation or modulation of pain.
CC -1- SUBCELLULAR LOCATION: Integral membrane protein.
CC -1- SIMILARITY: Belongs to family 1 of G-protein coupled receptors.
CC Was subfamily.

CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
CC or send an email to license@isb-sib.ch).
CC -----
DR EMBL, AF518238; AAC08310.1; -.
DR RGD; 738050; Mrgpra.
DR InterPro; IPR000276; GPCR_Rhodpsn.
DR Pfam; PF00001; 7tm 1; 1.
DR PROSITE; PS00237; G_PROTEIN_RECEP_F1_1; 1.
DR PROSITE; PS50262; G_PROTEIN_RECEP_F1_2; 1.
KW G-protein coupled receptor; Transmembrane.
FT DOMAIN 1 17 Extracellular (Potential).
FT TRANSMEM 18 38 Cytoplasmic (Potential).
FT DOMAIN 39 46 Cytoplasmic (Potential).
FT TRANSMEM 47 67 Extracellular (Potential).
FT DOMAIN 68 80 Extracellular (Potential).
FT TRANSMEM 81 101 Cytoplasmic (Potential).
FT DOMAIN 102 132 Cytoplasmic (Potential).
FT TRANSMEM 133 153 Extracellular (Potential).
FT DOMAIN 154 167 Extracellular (Potential).
FT TRANSMEM 168 188 Cytoplasmic (Potential).
FT DOMAIN 189 211 Cytoplasmic (Potential).
FT TRANSMEM 212 232 Extracellular (Potential).
FT DOMAIN 233 244 Extracellular (Potential).
FT TRANSMEM 245 265 Cytoplasmic (Potential).
FT DOMAIN 266 304 Cytoplasmic (Potential).
SQ SEQUENCE 304 AA; 34334 MW; 4A8204A0192E8B6 CRC64;

Query Match 4.5%; Score 15; DB 1; Length 304;
Best Local Similarity 100.0%; Pred. No. 6.7e-06;
Matches 15; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 124 GLSMLSAISTERCLS 138
Db 98 GLSMLSAISTERCLS 112

RESULT 13
MGA7_MOUSE STANDARD; PRT; 305 AA.
ID MGA7_MOUSE STANDARD; PRT; 305 AA.
AC Q912C5;
DT 01-OCT-2004 (Rel. 45, Created)
DT 01-OCT-2004 (Rel. 45, Last sequence update)
DT 01-OCT-2004 (Rel. 45, Last annotation update)
DE Mas-related G-protein coupled receptor member A7.
GN Name=Mrgpra7; Synonyms=Mrga7;
OS Mus musculus (Mouse).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
OX NCBI_TaxID=10090;
RN [1]
RP SEQUENCE FROM N.A., AND TISSUE SPECIFICITY.
RC STRAIN=129/SvJ;
RX MEDLINE=21435808; PubMed=11551509;
RA Dong X., Han S.-K., Zylka M.J., Simon M.I., Anderson D.J.;
RT "A diverse family of GPCRs expressed in specific subsets of
RT nociceptive sensory neurons."
RL Cell 106:619-632(2001).
CC -1- FUNCTION: Orphan receptor. May be a receptor for Rfam1de-family
CC neuropeptides such as NPY and NPAF, which are analgesic in vivo.
CC May regulate nociceptor function and/or development, including the
CC sensation or modulation of pain (By similarity).
CC -1- SUBCELLULAR LOCATION: Integral membrane protein.
CC -1- TISSUE SPECIFICITY: Expressed in a subset of sensory neurons that
CC includes nociceptors. Expressed in the subbase of nonpeptidergic
CC sensory neurons that are IB4(+) and VR1(-).
CC -1- SIMILARITY: Belongs to family 1 of G-protein coupled receptors.
CC Was subfamily.

CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
CC or send an email to license@isb-sib.ch).

DR EMBL: AY042197; AK91793.1; -
DR MGD; MG1:3033109; Mrgrpr8.
DR InterPro: IPR000276; GPCR_Rhodopsin.
DR Pfam; PF00001; 7tm_1; 1.
DR PRINTS; PR00237; GPCRHDOPS_N.
DR PROSITE; PS00237; G_PROTEIN_RECP_F1_1; 1.
DR PROSITE; PS0262; G_PROTEIN_RECP_F1_2; 1.
KW G-protein coupled receptor; Transmembrane.
FT DOMAIN 1 17 Extracellular (Potential).
FT TRANSMEM 18 38 1 (Potential).
FT DOMAIN 39 46 Cytoplasmic (Potential).
FT TRANSMEM 47 67 2 (Potential).
FT DOMAIN 68 81 Extracellular (Potential).
FT TRANSMEM 82 102 3 (Potential).
FT DOMAIN 103 129 Cytoplasmic (Potential).
FT TRANSMEM 130 150 4 (Potential).
FT DOMAIN 151 167 Extracellular (Potential).
FT TRANSMEM 168 188 5 (Potential).
FT DOMAIN 189 211 Cytoplasmic (Potential).
FT TRANSMEM 212 232 6 (Potential).
FT DOMAIN 233 244 Extracellular (Potential).
FT TRANSMEM 245 265 7 (Potential).
FT DOMAIN 266 305 Cytoplasmic (Potential).
SQ SEQUENCE 305 AA; 35034 MW; 254FDB3B9E6D39 CRC64;

Query Match 4.5%; Score 15; DB 1; Length 305;
Best Local Similarity 100.0%; Pred. No. 6.7e-06;
Matches 15; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 124 GLSMLSAISTERCLS 138
Db 98 GLSMLSAISTERCLS 112

RESULT 14
MGAB_MOUSE
ID MGAB_MOUSE STANDARD; PRT; 305 AA.
AC G91ZC4;
DT 01-OCT-2004 (Rel. 45, Created)
DT 01-OCT-2004 (Rel. 45, Last sequence update)
DT 01-OCT-2004 (Rel. 45, Last annotation update)
DE Mas-related G-protein coupled receptor member A8.
GN Name=Mrgrpr8; Synonyms=Mrga8;
OS Mus musculus (Mouse).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
OX NCBI_TaxID=10090;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=129/SVJ;
RX MEDLINE=21435808; PubMed=11551509;
RA Dong X., Han S.-K., Zylika M.J., Simon M.I., Anderson D.J.;
RT "A diverse family of GPCRs expressed in specific subsets of
RT nociceptive sensory neurons."
RL Cell 106:619-632(2001).

CC -1- FUNCTION: Orphan receptor. May be a receptor for Rfamde-family
CC neuropeptides such as NPPF and NPAF, which are analgesic in vivo.
CC May regulate nociceptor function and/or development, including the
CC sensation or modulation of pain (By similarity).
CC -1- SUBCELLULAR LOCATION: Integral membrane protein.
CC -1- TISSUE SPECIFICITY: Expressed in a subset of sensory neurons that
CC includes nociceptors. Expressed in the subnucleus of nonpeptidergic
CC sensory neurons that are IB4(+) and VR1(-).
CC -1- SIMILARITY: Belongs to family 1 of G-protein coupled receptors.
CC Mas subfamily.

CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
CC or send an email to license@isb-sib.ch).

DR EMBL: AY042198; AK91794.1; -
DR MGD; MG1:3033111; Mrgrpr8.
DR InterPro: IPR000276; GPCR_Rhodopsin.
DR Pfam; PF00001; 7tm_1; 1.
DR PRINTS; PR00237; GPCRHDOPS_N.
DR PROSITE; PS00237; G_PROTEIN_RECP_F1_1; 1.
DR PROSITE; PS0262; G_PROTEIN_RECP_F1_2; 1.
KW G-protein coupled receptor; Transmembrane.
FT DOMAIN 1 17 Extracellular (Potential).
FT TRANSMEM 18 38 1 (Potential).
FT DOMAIN 39 46 Cytoplasmic (Potential).
FT TRANSMEM 47 67 2 (Potential).
FT DOMAIN 68 85 Extracellular (Potential).
FT TRANSMEM 86 106 3 (Potential).
FT DOMAIN 107 129 Cytoplasmic (Potential).
FT TRANSMEM 130 150 4 (Potential).
FT DOMAIN 151 172 Extracellular (Potential).
FT TRANSMEM 173 193 5 (Potential).
FT DOMAIN 194 207 Cytoplasmic (Potential).
FT TRANSMEM 208 228 6 (Potential).
FT DOMAIN 229 243 Extracellular (Potential).
FT TRANSMEM 244 264 7 (Potential).
FT DOMAIN 265 305 Cytoplasmic (Potential).
FT CARBOHYD 151 151 N-linked (GlcNAc...) (Potential).
FT CARBOHYD 159 159 N-linked (GlcNAc...) (Potential).
SQ SEQUENCE 305 AA; 35004 MW; D6BB232DA58534F CRC64;

Query Match 4.5%; Score 15; DB 1; Length 305;
Best Local Similarity 100.0%; Pred. No. 6.7e-06;
Matches 15; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 124 GLSMLSAISTERCLS 138
Db 98 GLSMLSAISTERCLS 112

RESULT 15
MGAB_MOUSE
ID MGAB_MOUSE STANDARD; PRT; 313 AA.
AC G91WZ2;
DT 01-OCT-2004 (Rel. 45, Created)
DT 01-OCT-2004 (Rel. 45, Last sequence update)
DT 01-OCT-2004 (Rel. 45, Last annotation update)
DE Mas-related G-protein coupled receptor member A4 (RF-amide G protein-
DE coupled receptor).
GN Name=Mrgrpr4; Synonyms=Mrga4;
OS Mus musculus (Mouse).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
OX NCBI_TaxID=10090;
RN [1]
RP SEQUENCE FROM N.A.; SUBCELLULAR LOCATION, AND TISSUE SPECIFICITY.
RC STRAIN=C57BL/6; TISSUE=Dorsal root ganglion;
RX MEDLINE=21435808; PubMed=11551509;
RA Dong X., Han S.-K., Zylika M.J., Simon M.I., Anderson D.J.;
RT "A diverse family of GPCRs expressed in specific subsets of
RT nociceptive sensory neurons."
RL Cell 106:619-632(2001).

CC -1- FUNCTION: Orphan receptor. May be a receptor for Rfamde-family
CC neuropeptides such as NPPF and NPAF, which are analgesic in vivo.
CC May regulate nociceptor function and/or development, including the
CC sensation or modulation of pain.
CC -1- SUBCELLULAR LOCATION: Integral membrane protein.
CC -1- TISSUE SPECIFICITY: Expressed in a subset of sensory neurons that

```
CC includes nociceptors. Expressed in the subclass of nonpeptidergic
CC sensory neurons that are IB4(+) and VR1(-).
CC -1- SIMILARITY: Belongs to family 1 of G-protein coupled receptors.
CC Mas subfamily.
-----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
-----
CC
CC EMBL; AY042194; AAK91790.1; -.
CC MGD; MGI:3033100; Mrp4.
CC InterPro; IPR000276; GPCR_Rhodpsn.
CC Pfam; PF00001; 7tm.1; 1
CC PRINTS; PR00237; GFCRRHODPSN.
CC PROSITE; PS00237; G_PROTEIN_RECEP_F1_1; 1.
CC PROSITE; PS50262; G_PROTEIN_RECEP_F1_2; 1.
CC G-protein coupled receptor; Glycoprotein; Transmembrane.
CC DOMAIN 1 25 Extracellular (Potential).
CC TRANSMEM 26 46 1 (Potential).
CC DOMAIN 47 54 Cytoplasmic (Potential).
CC TRANSMEM 55 75 2 (Potential).
CC DOMAIN 76 93 Extracellular (Potential).
CC TRANSMEM 94 114 3 (Potential).
CC DOMAIN 115 137 Cytoplasmic (Potential).
CC TRANSMEM 138 158 4 (Potential).
CC DOMAIN 159 182 Extracellular (Potential).
CC TRANSMEM 183 203 5 (Potential).
CC DOMAIN 204 219 Cytoplasmic (Potential).
CC TRANSMEM 220 240 6 (Potential).
CC DOMAIN 241 255 Extracellular (Potential).
CC TRANSMEM 256 276 7 (Potential).
CC DOMAIN 277 313 Cytoplasmic (Potential).
CC CARBOHYD 10 10 N-linked (GlcNAc...) (Potential).
CC CARBOHYD 76 76 N-linked (GlcNAc...) (Potential).
CC SEQUENCE 313 AA; 35667 MW; 22F3AEC2F2F7127B CRC64;
```

```
Query Match 4.5%; Score 15; DB 1; Length 313;
Best Local Similarity 100.0%; Pred. No. 6.8e-06;
Matches 15; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
```

```
OY 124 GLSMLSAISTERCLS 138
Db 106 GLSMLSAISTERCLS 120
```

Search completed: October 27, 2004, 08:51:36
Job time : 197 secs

GenCore version 5.1.6
Copyright (c) 1993 - 2004 CompuGen Ltd.

OM nucleic - nucleic search, using sw model

Run on: October 29, 2004, 10:19:03 ; Search time 215 Seconds
(without alignments)

8655.094 Million cell updates/sec

Title: US-09-867-570-1
Perfect score: 2618
Sequence: 1 aacaatgcgcgcgaattcgcg.....gaaaaaaaaaaaaaaaaa 2618

Scoring table: IDENTITY NUC
Gap 10.0 , Gapext 1.0

Searched: 824507 seqs, 355394441 residues

Total number of hits satisfying chosen parameters: 1649014

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database :

- 1: Issued Patents NA:*
- 2: /cgn2_6/ptodata/1/ina/5A.COMB.seq:*
- 3: /cgn2_6/ptodata/1/ina/5B.COMB.seq:*
- 4: /cgn2_6/ptodata/1/ina/6A.COMB.seq:*
- 5: /cgn2_6/ptodata/1/ina/6B.COMB.seq:*
- 6: /cgn2_6/ptodata/1/ina/backfile1.seq:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	949.8	36.3	969	US-09-254-227A-4	Sequence 4, Appl1
2	938.6	35.9	969	US-09-254-227A-6	Sequence 6, Appl1
3	836.2	31.9	969	US-09-254-227A-12	Sequence 12, Appl1
4	833	31.8	969	US-09-254-227A-8	Sequence 8, Appl1
5	829.8	31.7	969	US-09-254-227A-14	Sequence 14, Appl1
6	826.6	31.6	969	US-09-254-227A-10	Sequence 10, Appl1
7	410.2	15.7	1011	US-09-254-227A-2	Sequence 2, Appl1
8	291	11.1	291	US-09-495-050A-164	Sequence 164, App
9	275	10.5	275	US-09-016-434-330	Sequence 330, App
C 10	146.4	5.6	148567	US-09-801-876B-3	Sequence 3, Appl1
C 11	146.4	5.6	148567	US-10-254-869-3	Sequence 3, Appl1
C 12	132.2	5.0	12047	US-09-032-461-1	Sequence 1, Appl1
C 13	132.2	5.0	12047	US-09-033-556-3	Sequence 3, Appl1
C 14	132.2	5.0	12047	US-09-474-699-11	Sequence 11, Appl1
C 15	132.2	5.0	12047	US-09-151-376-3	Sequence 3, Appl1
C 16	132.2	5.0	12047	US-09-814-351-11	Sequence 11, Appl1
C 17	126.8	4.7	118067	US-09-497-855A-32	Sequence 32, Appl1
C 18	123.4	4.7	11288	US-08-646-301A-1	Sequence 1, Appl1
C 19	123.4	4.7	11288	US-08-481-968A-4	Sequence 4, Appl1
C 20	123.4	4.7	11288	US-08-154-712B-4	Sequence 4, Appl1
C 21	123.4	4.7	11288	US-09-947-925A-4	Sequence 4, Appl1
C 22	123.4	4.7	15056	US-09-474-699-10	Sequence 10, Appl1
C 23	123.4	4.7	15056	US-09-814-351-14	Sequence 14, Appl1
C 24	122.8	4.7	14364	US-10-067-443-20	Sequence 20, Appl1
C 25	120.6	4.6	45716	US-08-965-048-5	Sequence 5, Appl1
C 26	120.6	4.6	45989	US-08-965-048-6	Sequence 6, Appl1
C 27	118.6	4.5	1388	US-09-016-434-1225	Sequence 1225, Ap

C 28	118.6	4.5	1388	5	PCT-US93-06251-26	Sequence 26, Appl1
C 29	114.4	4.4	3216	2	US-08-828-007-1	Sequence 1, Appl1
C 30	114.2	4.4	327	2	US-09-513-999C-25490	Sequence 25490, A
C 31	113.2	4.3	44453	3	US-09-146-053-5	Sequence 5, Appl1
C 32	112	4.3	505	3	US-09-227-357-73	Sequence 73, Appl1
C 33	111.6	4.3	1000	4	US-09-671-317-221	Sequence 221, App
C 34	111.6	4.3	1001	4	US-09-671-317-219	Sequence 219, App
C 35	111.6	4.3	1001	4	US-09-671-317-220	Sequence 220, App
C 36	111.6	4.3	1001	4	US-09-671-317-222	Sequence 222, App
C 37	110.8	4.2	2435	4	US-09-484-970B-134	Sequence 134, App
C 38	105.4	4.0	152331	3	US-09-128-155-16	Sequence 16, Appl1
C 39	105.4	4.0	176373	3	US-09-128-155-17	Sequence 17, Appl1
C 40	105	4.0	54945	4	US-09-967-669-10	Sequence 10, Appl1
C 41	104.4	4.0	152331	3	US-09-128-155-16	Sequence 16, Appl1
C 42	103.6	4.0	319608	4	US-09-539-333D-1	Sequence 1, Appl1
C 43	103.6	4.0	319608	4	US-09-679-409-1	Sequence 1, Appl1
C 44	103.4	3.9	438	2	US-09-621-976-18282	Sequence 18282, A
C 45	102.6	3.9	246240	2	US-08-724-394A-20	Sequence 20, Appl1

ALIGNMENTS

RESULT 1						
US-09-254-227A-4						
; Sequence 4, Application US/09254227A						
; Patent No. 6696257						
; GENERAL INFORMATION:						
; APPLICANT: Ahmad, Sultan						
; APPLICANT: Banville, Denis						
; APPLICANT: Fortin, Yves						
; APPLICANT: Lembo, Paola						
; APPLICANT: O'Donnell, Dajan						
; APPLICANT: Shi-Hsiang, Shen						
; TITLE OF INVENTION: G Protein-Coupled Receptors from the Rat and Human						
; FILE REFERENCE: 81823/268117						
; CURRENT APPLICATION NUMBER: US/09/254,227A						
; CURRENT FILING DATE: 1999-03-03						
; NUMBER OF SEQ ID NOS: 22						
; SOFTWARE: PatentIn version 3.0						
; SEQ ID NO 4						
; LENGTH: 969						
; TYPE: DNA						
; ORGANISM: Homo sapiens						
US-09-254-227A-4						
Query Match 36.3%; Score 949.8; DB 4; Length 969;						
Best Local Similarity 98.8%; Pred. No. 2,2e-263;						
Matches 957; Conservative 0; Mismatches 12; Indels 0; Gaps 0;						
QY	492	ATGATTCACACCATCCAGTCTTGGGTACAGAACTGACACCAATCAACGACGAGAG	551			
DB	1	ATGATTCACACCATCCAGTCTTGGGTACAGAACTGACACCAATCAACGACGAGAG	60			
QY	552	ACTGCTGTACACGACGACCTGAGCTTACGCGGCTGACGTCATGTTCCCTTGTG	611			
DB	61	ACTGCTGTACACGACGACCTGAGCTTACGCGGCTGACGTCATGTTCCCTTGTG	120			
QY	612	CGCGTACAGAAACGGGTTGTGCTGTGCTCTGCGCTTCCGATCGCAGAAAGCT	671			
DB	121	CGCGTACAGAAACGGGTTGTGCTGTGCTCTGCGCTTCCGATCGCAGAAAGCT	180			
QY	672	GTCCTCATTCACATCCCTCAACCTGTGCGCGCACTTCCCTTCTTAAGGCGCACTT	731			
DB	181	GTCCTCATTCACATCCCTCAACCTGTGCGCGCACTTCCCTTCTTAAGGCGCACTT	240			
QY	732	ATATGTTGCGCGTACCGCTCATATATCGCCATCCATCTCCAAATCTCGATCTCT	791			
DB	241	ATATGTTGCGCGTACCGCTCATATATCGCCATCCATCTCCAAATCTCGATCTCT	300			
QY	792	GTGATGACCTTTCCCTTATATAGCCTTAAGCATGCTGAGCGCCATCAGCAGAGCGC	851			
DB	301	GTGATGACCTTTCCCTTATATAGCCTTAAGCATGCTGAGCGCCATCAGCAGAGCGC	360			

; APPLICANT: Lembo, Paola
 ; APPLICANT: O'Donnell, Dajan
 ; APPLICANT: Shi-Heiang, Shen
 ; TITLE OF INVENTION: G Protein-Coupled Receptors from the Rat and Human
 ; FILE REFERENCE: 81823/268117
 ; CURRENT APPLICATION NUMBER: US/09/254,227A
 ; CURRENT FILING DATE: 1999-03-03
 ; NUMBER OF SEQ ID NOS: 22
 ; SOFTWARE: PatentIn version 3.0
 ; SEQ ID NO: 10
 ; LENGTH: 969
 ; TYPE: DNA
 ; ORGANISM: Homo sapiens
 ; US-09-254-227A-10

Query Match 31.6%; Score 826.6; DB 4; Length 969;
 Best Local Similarity 90.8%; Pred. No. 7.2e-228;
 Matches 880; Conservative 0; Mismatches 89; Indels 0; Gaps 0;

QY 492 ATGATTCACACATCCAGTCTTGGTACAGAACTGACACCAATCAACGAGCGTAGAG 551
 DB 1 ATGATTCACACGCTCTCACTTGAACAAGAAATGACCAATCAACGAGACTGAGAG 60
 QY 552 ACTCTTGCTACACAGACACCTGAGCTTCAAGGAGCTGACGTGATCGTTTCCCTTGTG 611
 DB 61 ACTCTTGCTACACAGACACCTGAGCTTCAAGGAGCTGACGTGATCGTTTCCCTTGTG 120
 QY 612 GCGCTGACAGAAACGCGGTGTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 671
 DB 121 GCGCTGACAGAAACGCGGTGTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 180
 QY 672 GTCTCCATCTACACCTCAACCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 731
 DB 181 TTCTCCATCTACACCTCAACCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 240
 QY 732 ATATGTCGCGGTACCGCTCATCAATATCCGCAATCCCAATCCCAATCCCAATCCCAAT 791
 DB 241 ATATATTCCTGTTAAGCTTATCAATATCCGCAATCCCAATCCCAATCCCAATCCCAAT 300
 QY 792 GTGATGACCTTTCCTTATTAAGGCTTAAGCAATGCTGAGCGCAATCAGACGAGCGC 851
 DB 301 GTGATGATGTTTCTTCTTCTTCTGAGCGCTGAGCTTCTGAGTGCCTGAGACGAGCGC 360
 QY 852 TGCTGTCACATCCGTCGTCATCTGTCATCTGTCATCTGTCATCTGTCATCTGTCATCTG 911
 DB 361 TGCTGTCGCTGTCGTCATCTGTCATCTGTCATCTGTCATCTGTCATCTGTCATCTGTC 420
 QY 912 GTCATGTGTCTGCTCTGAGGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 971
 DB 421 GTGATGTGTCTGCTCTGAGGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 480
 QY 972 TGTGATCTCTGTTAAGTGTGCTGATTTCTGTTGTGTGTAAGAGTCAGATTTCAATACA 1031
 DB 481 TGTGATCTCTGTTAAGTGTGCTGATTTCTGTTGTGTGTAAGAGTCAGATTTCAATACA 540
 QY 1032 ATGCGGTGCTGCTGTTTATATGTGTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 1091
 DB 541 GTGCGGTGCTGATTTTATATGTGTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 600
 QY 1092 AGATTTCTGTTGATCCCGAAGATGCGGCTGACAGAGGCTGATCGTACATCCCTCTC 1151
 DB 601 AGATTTCTGTTGATCCCGAAGATGCGGCTGACAGAGGCTGATCGTACATCCCTCTC 660
 QY 1152 ACAGTGTGTCTTCTCTCTGTCGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 1211
 DB 661 ACAGTGTGTCTTCTCTCTGTCGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 720
 QY 1212 AGATTCACCTGATTTGAAGTCTTATTTGTCATGTCATGTCATGTCATGTCATGTCATGTC 1271
 DB 721 TGAATTCACCTGATTTGAAGTCTTATTTGTCATGTCATGTCATGTCATGTCATGTCATGTC 780
 QY 1272 TCCGCTTAAACAGAGTGCACCCCAATCAATTAATCTTCTGTCGCTGCTGTCGCTGTCGCT 1331

DB 781 TCCGCTTAAACAGAGTGCACCCCAATCAATTAATCTTCTGTCGCTGCTGTCGCTGTCGCT 840
 QY 1332 CGTCAAAATAGGAGAGACCTGAAGCTGTTCTGACAGAGGCTGTGACAGACCCCTGAG 1391
 DB 841 CGTCAAAATAGGAGAGACCTGAAGCTGTTCTGACAGAGGCTGTGACAGAGCCCTGAG 900
 QY 1392 GTGATGAAGTGAAGGTGCTTCTGACAGAAACCTGAGCTGTGCGGAAGAGATTG 1451
 DB 901 GTGATGAAGTGAAGGTGAGGCTTCTGAGGAATCTGAGAGCTGTGCGGAAGAGATTG 960
 QY 1452 GAGCAGTGA 1460
 DB 961 GAGCAGTGA 969

RESULT 7
 US-09-254-227A-2
 ; Sequence 2, Application US/09254227A
 ; Patent No. 6696257
 ; GENERAL INFORMATION:
 ; APPLICANT: Ahmad, Sultan
 ; APPLICANT: Banville, Denis
 ; APPLICANT: Fortin, Yves
 ; APPLICANT: Lembo, Paola
 ; APPLICANT: O'Donnell, Dajan
 ; APPLICANT: Shi-Heiang, Shen
 ; TITLE OF INVENTION: G Protein-Coupled Receptors from the Rat and Human
 ; FILE REFERENCE: 81823/268117
 ; CURRENT APPLICATION NUMBER: US/09/254,227A
 ; CURRENT FILING DATE: 1999-03-03
 ; NUMBER OF SEQ ID NOS: 22
 ; SOFTWARE: PatentIn version 3.0
 ; SEQ ID NO: 2
 ; LENGTH: 1011
 ; TYPE: DNA
 ; ORGANISM: rat
 ; US-09-254-227A-2

Query Match 15.7%; Score 410.2; DB 4; Length 1011;
 Best Local Similarity 65.6%; Pred. No. 8.4e-108;
 Matches 655; Conservative 0; Mismatches 323; Indels 21; Gaps 3;

QY 466 GGGTCATCAAGCTGGGTTTCTGAGCATGATTCACATCCAGTCTTGGGTACAGAAC 525
 DB 17 GGGACATCACTGGAAGATTGTGAGCATGATTCACATCTCTATCCCTAGTACAGAT 76
 QY 526 TGACCAATCAACGAGCGTGAAGGAGCTCTTGTCTACAGACACCTGAGCTTCAAG 585
 DB 77 CTACACACCTGAATTAACATGCTCA---TCCAGTTGCAAGGCCAATCTCACTCTGCT 133
 QY 586 GGGTCAGTGTGTTTCTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 645
 DB 134 TCCGTCGCCCATCATCACTGCTGATGAGGAGAAACACATGTTACTTGGCTCT 193
 QY 646 TGCGCTGCGCATGCGAGAGACGCTGCTCATCTCATCTCAATCTGATCTGCTGCTGCTGCT 705
 DB 194 TGGAATTCGCAATCGAGAGAAACATCTCAAGTCAAGTCAAGTCAAGTCAAGTCAAGTCA 253
 QY 706 ACTTCTCTTCTTGAAGGCGCATATTAATGTTGCGCTTACGCT----- 751
 DB 254 ACTCTCTTCTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 313
 QY 752 -CATCAATATCCGATCCCATCTCCAAATCTCAAGTCTGATGATGATGATGATGATGATGAT 810
 DB 314 GCATCTATGCGCATTAATTAAGCAAGAAATCTTAAGCAATGATGATGATGATGATGATGAT 373
 QY 811 TTATAGGCTTAAGCATGCTGAGCGCATTCAGACAGAGCGCTGCTGCTGCTGCTGCTGCTGCT 870
 DB 374 TCTAGGCTGAGCATCTCAAGTCTATCAAGACAGAGCGCTGCTGCTGCTGCTGCTGCTGCT 433
 QY 871 CCATCTGATCACTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 930
 DB 434 CAATCTGATCACTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 493

REGISTRATION NUMBER: 40,130
REFERENCE/DOCKET NUMBER: 34802-20003.21
TELECOMMUNICATION INFORMATION:
TELEPHONE: 415-813-5600
TELEFAX: 415-494-0792
TELEX: 706141
INFORMATION FOR SEQ ID NO: 1:
SEQUENCE CHARACTERISTICS:
LENGTH: 12047 base pairs
TYPE: nucleic acid
STRANDEDNESS: single
TOPOLOGY: linear
US-09-022-461-1

Query Match 5.0%; Score 132.2; DB 2; Length 12047;
Best Local Similarity 75.9%; Pred. No. 6.2e-27;
Matches 189; Conservative 0; Mismatches 58; Indels 2; Gaps 2;
QY 28 TGAATCTAGTGTGTTAAAGGCTGTAGACACTCTCCCTCTCTTACTCTGCTCTCA 87
DB 5320 TGAGATCTGTCATTTAAAGTGCTGCGCCCTCCCTCTCTCTGCTCTCTACTG 5379
QY 88 CCATGTGAGACGCTCGCTCCCTCCCTTTCCTTACAGAGATTGAGAGCTTCTGTAGGCC 147
DB 5380 CCATGTAGATACCT-GCTCCGCTTTGCTTCTACCAATAGTAAAGCCCTGTAGGCC 5438
QY 148 TCCCGAAGAGAGAGAGCTGCTATGCTTCTGTACAGCTGTAGAGCTATTAGCCAGTTAA 207
DB 5439 TCCCGAAG 5498
QY 208 ACCCATCTCTCATTAATTTCCAGCTCAGAGATTCTTTTGAACAATTGAGATGA 267
DB 5499 ACCCTCTTCTGTATATA-TTACAGCTGTGAGTATCTCTTACAGAGAGTGTAGAACG 5557
QY 268 ACTAATACA 276
DB 5558 ACTAATACA 5566

RESULT 13
US-09-033-556-3
Sequence 3, Application US/09033556
Patent No. 6432700
GENERAL INFORMATION:
APPLICANT: Henderson, Daniel R.
APPLICANT: Yu, De Chao
TITLE OF INVENTION: ADENOVIRUS VECTORS CONTAINING
TITLE OF INVENTION: HETEROLOGOUS TRANSCRIPTION REGULATORY ELEMENTS AND METHODS
TITLE OF INVENTION: OF USING SAME
NUMBER OF SEQUENCES: 41
CORRESPONDENCE ADDRESS:
ADDRESSEE: MORRISON & FOERSTER
STREET: 755 PAGE MILL ROAD
CITY: Palo Alto
STATE: CA
COUNTRY: USA
ZIP: 94304-1018
COMPUTER READABLE FORM:
MEDIUM TYPE: Diskette
COMPUTER: IBM Compatible
OPERATING SYSTEM: Windows
SOFTWARE: FastSeq for Windows Version 2.0b
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/09/033,556
FILING DATE:
CLASSIFICATION:
PRIOR APPLICATION DATA:
APPLICATION NUMBER:
FILING DATE:
ATTORNEY/AGENT INFORMATION:
NAME: Catherine, Polizzi M
REGISTRATION NUMBER: 40,130
REFERENCE/DOCKET NUMBER: 34802-20010.00

TELECOMMUNICATION INFORMATION:
TELEPHONE: 650-813-5600
TELEFAX: 650-494-0792
TELEX: 706141
INFORMATION FOR SEQ ID NO: 3:
SEQUENCE CHARACTERISTICS:
LENGTH: 12047 base pairs
TYPE: nucleic acid
STRANDEDNESS: single
TOPOLOGY: linear
US-09-033-556-3

Query Match 5.0%; Score 132.2; DB 4; Length 12047;
Best Local Similarity 75.9%; Pred. No. 6.2e-27;
Matches 189; Conservative 0; Mismatches 58; Indels 2; Gaps 2;
QY 28 TGAATCTAGTGTGTTAAAGGCTGTAGACACTCTCCCTCTCTTACTCTGCTCTCA 87
DB 5320 TGAGATCTGTCATTTAAAGTGCTGCGCCCTCCCTCTCTCTGCTCTCTACTG 5379
QY 88 CCATGTGAGACGCTCGCTCCCTCCCTTTCCTTACAGAGATTGAGAGCTTCTGTAGGCC 147
DB 5380 CCATGTAGATACCT-GCTCCGCTTTGCTTCTACCAATAGTAAAGCCCTGTAGGCC 5438
QY 148 TCCCGAAGAGAGAGAGCTGCTATGCTTCTGTACAGCTGTAGAGCTATTAGCCAGTTAA 207
DB 5439 TCCCGAAG 5498
QY 208 ACCCATCTCTCATTAATTTCCAGCTCAGAGATTCTTTTGAACAATTGAGATGA 267
DB 5499 ACCCTCTTCTGTATATA-TTACAGCTGTGAGTATCTCTTACAGAGAGTGTAGAACG 5557
QY 268 ACTAATACA 276
DB 5558 ACTAATACA 5566

RESULT 14
US-09-474-699-11
Sequence 11, Application US/09474699
Patent No. 6495130
GENERAL INFORMATION:
APPLICANT: Henderson, Daniel R.
APPLICANT: Yu, De Chao
TITLE OF INVENTION: TARGET CELL-SPECIFIC ADENOVIRAL VECTORS
TITLE OF INVENTION: CONTAINING E3 AND METHODS OF USE THEREOF
FILE REFERENCE: 348022001300
CURRENT APPLICATION NUMBER: US/09/474,699
PRIOR FILING DATE: 1999-12-29
PRIOR APPLICATION NUMBER: 60/114,262
NUMBER OF SEQ ID NOS: 23
SOFTWARE: FastSeq for Windows Version 4.0
SEQ ID NO 11
LENGTH: 12047
TYPE: DNA
ORGANISM: Homo Sapien
US-09-474-699-11

Query Match 5.0%; Score 132.2; DB 4; Length 12047;
Best Local Similarity 75.9%; Pred. No. 6.2e-27;
Matches 189; Conservative 0; Mismatches 58; Indels 2; Gaps 2;
QY 28 TGAATCTAGTGTGTTAAAGGCTGTAGACACTCTCCCTCTCTTACTCTGCTCTCA 87
DB 5320 TGAGATCTGTCATTTAAAGTGCTGCGCCCTCCCTCTCTCTGCTCTCTACTG 5379
QY 88 CCATGTGAGACGCTCGCTCCCTCCCTTTCCTTACAGAGATTGAGAGCTTCTGTAGGCC 147
DB 5380 CCATGTAGATACCT-GCTCCGCTTTGCTTCTACCAATAGTAAAGCCCTGTAGGCC 5438
QY 148 TCCCGAAGAGAGAGAGAGCTGCTATGCTTCTGTACAGCTGTAGAGCTATTAGCCAGTTAA 207

Db 5439 TCCCGAAGCAGATGCCAGCTTCTGTACAGCTTCAGAACCATCAGCCAAATTAA 5498
QY 208 ACCATTTCCTTCAATTAATTTCCAGTCTCAGGATTTCTTTAGCAATTTGAGATGA 267
Db 5499 ACCTCTTTCTGTATAAA-TTACCAAGCTTGAGATCTCTTAACAGCAGTGTAGAACGG 5557
QY 268 ACTAATACA 276
Db 5558 ACTAATACA 5566

RESULT 15

US-09-151-376-3
; Sequence 3, Application US/09151376
; Patent No. 6676935
; GENERAL INFORMATION:
; APPLICANT: Henderson, D.R.
; APPLICANT: Schuur, B.R.
; TITLE OF INVENTION: TISSUE SPECIFIC VIRAL VECTORS
; FILE REFERENCE: 34802200221
; CURRENT APPLICATION NUMBER: US/09/151,376
; EARLIER FILING DATE: 1998-09-10
; EARLIER APPLICATION NUMBER: 08/669,753
; EARLIER FILING DATE: 1996-06-26
; EARLIER APPLICATION NUMBER: 08/495,034
; EARLIER FILING DATE: 1995-06-27
; NUMBER OF SEQ ID NOS: 71
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 3
; LENGTH: 12047
; TYPE: DNA
; ORGANISM: Homo sapiens
US-09-151-376-3

Query Match 5.0%; Score 132.2; DB 4; Length 12047;
Best Local Similarity 75.9%; Pred. No. 6.2e-27;
Matches 189; Conservative 0; Mismatches 58; Indels 2; Gaps 2;

QY 28 TGAATCTAGTTGTTTAAAGCGGTAGACCTCCCTCTCTTACTCTGCTCTCA 87
Db 5320 TGAGATCTGTCAATTAAAGGTGTGGCCCTCCCTCTCTGTGTCCTACTG 5379
QY 88 CCATGTGAGAGCGCTCGCTCCCTTGCCTTTCACCAAGGATTGAAAGCTTCTGAGGCC 147
Db 5380 CCATGTAGATTAAGT-CTCTGCTTGTGCTTCTACCATAGTAAAGCCCCCTGAGGCC 5438
QY 148 TCCCGAAGCAGAGCTGCTATCTTCTGTACAGTCTGTAGAGCTATTAGCCAGTTAA 207
Db 5439 TCCCGAAGCAGATGCACCATGCTTCTGTACAGCTGCAGAACCATCAGCCAAATTAA 5498
QY 208 ACCCATTTCTTCAATTAATTTCCAGTCTCAGGTATTTCTTTAGCAATTTGAAATGA 267
Db 5499 ACCTCTTTCTGTATAAA-TTACCAAGCTTGAGATCTCTTAACAGCAGTGTAGAACGG 5557
QY 268 ACTAATACA 276
Db 5558 ACTAATACA 5566

Search completed: October 29, 2004, 16:11:07
Job time : 220 secs

This Page Blank (uspto)

GenCore version 5.1.6
Copyright (c) 1993 - 2004 CompuGen Ltd.

OM nucleic - nucleic search, using SW model

Run on: October 29, 2004, 13:55:34 : Search time 1218 Seconds
(without alignments)
11021.397 Million cell updates/sec

Title: US-09-867-570-1

Perfect score: 2618
Sequence: 1 aacaaatggccgcgaattcg.....gaaaaaaaaaaaaaaaaa 2618

Scoring table: IDENTITY_NTC
Gapop 10.0 , Gapext 1.0

Searched: 3413475 seqs, 256380928 residues

Total number of hits satisfying chosen parameters: 6826950

Minimum DB seq length: 0

Maximum DB seq length: 200000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database :

Published Applications NA:*
1: /cgn2_6/ptodata/2/pubpna/US07_PUBCOMB.seq:*
2: /cgn2_6/ptodata/2/pubpna/US06_NEW_PUB.seq:*
3: /cgn2_6/ptodata/2/pubpna/US06_PUBCOMB.seq:*
4: /cgn2_6/ptodata/2/pubpna/US07_NEW_PUB.seq:*
5: /cgn2_6/ptodata/2/pubpna/US07_PUBCOMB.seq:*
6: /cgn2_6/ptodata/2/pubpna/US08_NEW_PUB.seq:*
7: /cgn2_6/ptodata/2/pubpna/US08_PUBCOMB.seq:*
8: /cgn2_6/ptodata/2/pubpna/US09_NEW_PUBCOMB.seq:*
9: /cgn2_6/ptodata/2/pubpna/US09_PUBCOMB.seq:*
10: /cgn2_6/ptodata/2/pubpna/US09C_PUBCOMB.seq:*
11: /cgn2_6/ptodata/2/pubpna/US09C_NEW_PUB.seq:*
12: /cgn2_6/ptodata/2/pubpna/US10_PUBCOMB.seq:*
13: /cgn2_6/ptodata/2/pubpna/US10C_PUBCOMB.seq:*
14: /cgn2_6/ptodata/2/pubpna/US10C_PUBCOMB.seq:*
15: /cgn2_6/ptodata/2/pubpna/US10C_PUBCOMB.seq:*
16: /cgn2_6/ptodata/2/pubpna/US10C_PUBCOMB.seq:*
17: /cgn2_6/ptodata/2/pubpna/US10C_PUBCOMB.seq:*
18: /cgn2_6/ptodata/2/pubpna/US11_NEW_PUB.seq:*
19: /cgn2_6/ptodata/2/pubpna/US11_NEW_PUBCOMB.seq:*
20: /cgn2_6/ptodata/2/pubpna/US60_NEW_PUBCOMB.seq:*
21: /cgn2_6/ptodata/2/pubpna/US60_PUBCOMB.seq:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	2618	100.0	2618	11	US-09-867-570-1
2	1529.6	58.4	1997	14	US-10-219-834-7
3	1483.8	56.7	2040	14	US-10-183-116-15
4	1192.6	45.6	1369	15	US-10-292-798-1273
5	1104	42.2	8622	11	US-09-867-570-3
6	1094.2	41.8	1400	14	US-10-183-116-30
7	1094.2	41.8	1400	15	US-10-225-567A-673
8	1053	40.2	1370	15	US-10-017-161-1599
9	1040.6	39.7	1369	15	US-10-017-161-1055
10	1040.6	39.7	1369	15	US-10-292-798-897
11	1013.6	38.7	1604	14	US-10-183-116-32
12	1013.6	38.7	1604	15	US-10-225-567A-688

13	1012.2	38.7	1369	15	US-10-292-798-1041	Sequence 1041, Ap
14	969	37.0	969	9	US-09-995-225-19	Sequence 19, Appl
15	969	37.0	969	10	US-09-995-225-19	Sequence 19, Appl
16	965.8	36.9	969	15	US-10-401-397A-1	Sequence 1, Appl
17	964.2	36.8	969	16	US-10-391-074-1	Sequence 1, Appl
18	856.2	32.7	997	16	US-10-072-012-111	Sequence 171, App
19	845.8	32.3	909	5	US-10-237-467-9	Sequence 9, Appl
20	833	31.8	969	15	US-09-995-225-17	Sequence 17, Appl
21	833	31.8	969	10	US-09-995-225-17	Sequence 17, Appl
22	833	31.8	969	15	US-10-237-467-3	Sequence 3, Appl
23	829.8	31.7	969	15	US-10-079-384-3	Sequence 3, Appl
24	829.8	31.7	969	15	US-10-240-998-3	Sequence 19, Appl
25	829.8	31.7	969	15	US-10-321-807-19	Sequence 19, Appl
26	829.8	31.7	969	15	US-10-237-467-11	Sequence 43, Appl
27	829.8	31.7	969	15	US-10-343-650A-43	Sequence 19, Appl
28	829.8	31.7	969	17	US-10-321-807-19	Sequence 19, Appl
29	829.8	31.7	969	17	US-10-314-048A-19	Sequence 173, App
30	823.2	31.4	1005	16	US-10-072-012-113	Sequence 177, App
31	822.6	30.3	966	15	US-10-072-012-117	Sequence 9, Appl
32	792.2	30.3	966	15	US-10-240-998-9	Sequence 169, App
33	787.4	30.1	994	15	US-10-072-012-169	Sequence 6, Appl
34	769.6	29.4	966	15	US-10-240-998-6	Sequence 239, App
35	755.4	28.9	769	15	US-10-101-510-239	Sequence 1077, Ap
36	740.2	28.3	955	15	US-10-292-798-1077	Sequence 1269, Ap
37	713.8	27.3	1394	15	US-10-292-798-1269	Sequence 1311, Ap
38	687.8	26.3	955	15	US-10-017-161-1311	Sequence 175, App
39	628.8	24.0	1030	16	US-10-072-012-175	Sequence 7, Appl
40	618.8	23.6	961	15	US-10-240-998-7	Sequence 11, Appl
41	608.8	23.3	1163	9	US-09-750-373-11	Sequence 11, Appl
42	608.8	23.3	1300	14	US-10-183-116-17	Sequence 648, App
43	608.8	23.3	1300	15	US-10-225-567A-648	Sequence 1053, Ap
44	608.8	23.3	1393	15	US-10-017-161-1053	Sequence 895, App
45	608.8	23.3	1393	15	US-10-292-798-895	

ALIGNMENTS

RESULT 1
US-09-867-570-1
; Sequence 1, Application US/09867570
; Publication No. US20040076951A1
; GENERAL INFORMATION:
; APPLICANT: WEI, Ming-Hui et al.
; TITLE OF INVENTION: ISOLATED HUMAN G-PROTEIN COUPLED
; TITLE OF INVENTION: RECEPTORS, NUCLEIC ACID MOLECULES ENCODING HUMAN GPCR
; FILE REFERENCE: CL000900-CIP
; CURRENT APPLICATION NUMBER: US/09/867,570
; CURRENT FILING DATE: 2001-05-31
; PRIOR APPLICATION NUMBER: 09/695,045
; PRIOR FILING DATE: 2000-10-25
; NUMBER OF SEQ ID NOS: 4
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 1
; LENGTH: 2618
; TYPE: DNA
; ORGANISM: Human
; US-09-867-570-1

Query Match 100.0%; Score 2618; DB 11; Length 2618;
Best Local Similarity 100.0%; Pred. No. 0;
Matches 2618; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY	1	AACAAATGGCCGCAATTCGCGACGAGATGAATCTAGTTGTTTAAAGCGTGAACCT	60
DB	1	AACAAATGGCCGCAATTCGCGACGAGATGAATCTAGTTGTTTAAAGCGTGAACCT	60
QY	61	CCCTCCCTCTCTCTTACTCTCTGCTCTCAACATGGAAGCGTCCGCCCTTGGCTTT	120
DB	61	CCCTCCCTCTCTCTTACTCTCTGCTCTCAACATGGAAGCGTCCGCCCTTGGCTTT	120
QY	121	CACCAAGATGGAAGCTTCTGAGCGCTCCCGACGAAGCAAGAGCTGTATGCTTCTTGA	180

121 CACGAGATTGGAAGCTTCTGAGGCCCTCCAGAGCAGAAAGCTGATCTTCTGTA 180
181 CAGCTGTAGAGCATTAAGCAGATTAAACCAATTTCCCTCAATAATTTCCAGTCTGAG 240
181 CAGCTGTAGAGCATTAAGCAGATTAAACCAATTTCCCTCAATAATTTCCAGTCTGAG 240
241 TATTTCTTTTGAAGAAATTTGAAGTGAATGAATGAATGAATGAATGAATGAATGAATGA 300
241 TATTTCTTTTGAAGAAATTTGAAGTGAATGAATGAATGAATGAATGAATGAATGAATGA 300
301 ATCCCAAGTGTCTTCTGCTGTCTTCCAGTCTCTGCTGTCTCTCCAGTGTCTCAAT 360
301 ATCCCAAGTGTCTTCTGCTGTCTTCCAGTCTCTGCTGTCTCTCCAGTGTCTCAAT 360
361 TCCACCGAAGAAACGAAATTAAGAAATCCCAATGATGATGATGATGATGATGATGATGAT 420
361 TCCACCGAAGAAACGAAATTAAGAAATCCCAATGATGATGATGATGATGATGATGATGAT 420
421 TGGGATGTCAAAACAGATTAAGAAATGAAGAAATGAAGAAATGAAGAAATGAAGAAATGA 480
421 TGGGATGTCAAAACAGATTAAGAAATGAAGAAATGAAGAAATGAAGAAATGAAGAAATGA 480
481 GGTTCGTGAGATGATGATTAACATCCAGTCTTGGATGATGATGATGATGATGATGATGAT 540
481 GGTTCGTGAGATGATGATTAACATCCAGTCTTGGATGATGATGATGATGATGATGATGAT 540
541 GAGTGTGAGAGAGCTCTTGTCTCAAGAGAGAGCTGATGATGATGATGATGATGATGATGAT 600
541 GAGTGTGAGAGAGCTCTTGTCTCAAGAGAGAGCTGATGATGATGATGATGATGATGATGAT 600
601 TTTTCCCTGTGAGCTGTGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 660
601 TTTTCCCTGTGAGCTGTGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 660
661 GCGAGAAAGCTGTCTCAATCTCAATCTCAATCTCAATCTCAATCTCAATCTCAATCTCAAT 720
661 GCGAGAAAGCTGTCTCAATCTCAATCTCAATCTCAATCTCAATCTCAATCTCAATCTCAAT 720
721 GCGGCGCAATTAATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 780
721 GCGGCGCAATTAATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 780
781 TCCCTGATCTGTGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 840
781 TCCCTGATCTGTGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 840
841 GCAACGAGCGCTGTCTCAATCTCAATCTCAATCTCAATCTCAATCTCAATCTCAATCTCAAT 900
841 GCAACGAGCGCTGTCTCAATCTCAATCTCAATCTCAATCTCAATCTCAATCTCAATCTCAAT 900
901 ACCGTCATGCGTGTGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 960
901 ACCGTCATGCGTGTGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 960
961 AGTGTGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 1020
961 AGTGTGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 1020
1021 ATTTTCAATTAAGAAATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 1080
1021 ATTTTCAATTAAGAAATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 1080
1081 TCCGTCGTGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 1140
1081 TCCGTCGTGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 1140
1141 CCAATCTCTCTCAAGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 1200
1141 CCAATCTCTCTCAAGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 1200
1201 CCGTGTTCAGAGATCACTGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 1260
1201 CCGTGTTCAGAGATCACTGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 1260

1201 CCGTGTTCAGAGATCACTGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 1260
1261 CCAATTTCTCTGTCCGCTCTTAACAGAGATGATGATGATGATGATGATGATGATGATGATGAT 1320
1261 CCAATTTCTCTGTCCGCTCTTAACAGAGATGATGATGATGATGATGATGATGATGATGATGAT 1320
1321 CTTTAAAGCAGCTGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 1380
1321 CTTTAAAGCAGCTGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 1380
1381 CTTTAAAGCAGCTGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 1440
1381 CTTTAAAGCAGCTGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 1440
1441 GAGCAGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 1500
1441 GAGCAGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 1500
1501 TGTGCTGCTGCACTGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 1560
1501 TGTGCTGCTGCACTGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 1560
1561 TCTGAGTGTGCTGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 1620
1561 TCTGAGTGTGCTGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 1620
1621 CACCATGAGAAAGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 1680
1621 CACCATGAGAAAGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 1680
1681 ACATTTTCTGCTGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 1740
1681 ACATTTTCTGCTGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 1740
1741 ATTTTAAAGAAAGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 1800
1741 ATTTTAAAGAAAGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 1800
1801 CCTGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 1860
1801 CCTGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 1860
1861 TCAATGTTCTGCTGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 1920
1861 TCAATGTTCTGCTGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 1920
1921 GAGTGTGCTGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 1980
1921 GAGTGTGCTGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 1980
1981 GAGAGAGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 2040
1981 GAGAGAGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 2040
2041 CTGTGATTTTGTCAACAGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 2100
2041 CTGTGATTTTGTCAACAGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 2100
2101 GCGCTTCTGCTGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 2160
2101 GCGCTTCTGCTGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 2160
2161 TGAGAGAGCTGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 2220
2161 TGAGAGAGCTGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 2220
2221 ATTTTAAAGTGAATTAATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 2280
2221 ATTTTAAAGTGAATTAATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 2280
2281 GTAGAGAAATTTCTCAATCTGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 2340
2281 GTAGAGAAATTTCTCAATCTGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 2340

Db	1453	TGCTGTCCACAACCTTAACCTCTTCTTAACTGTTCTCCACTGATGATGATCAAAA	15112
Qy	1818	GGAAATCTCCCTTAATATCTGTGCAGACATATGTCCTCCCGGAAAATCAGTCCCTTAATG	18777
Db	1513	GGAAATCTCCCTTAATATCTGTGCAGACATATGTCCTCCCGGAAAATCAGTCCCTTAATG	15722
Qy	1878	ACTGAGGACATTAACGACAGTGGAAAGCTCAATTCTTAATAAGTAGGTTCTGCTACCTCTA	19373
Db	1573	ACTGAGGACATTAACGACAGTGGAAAGCTCAATTCTTAATAAGTAGGTTCTGCTACCTCTA	16322
Qy	1938	AATTCATTTGAATTCTCAGATATTAAGCAAAATTAATGACTTTAGAGAGATTTCCCTT	19977
Db	1633	AATTCATTTGAATTCTCAGATATTAAGCAAAATTAATGACTTTAGAGAGATTTCCCTT	16922
Qy	1998	CATAAAACAGCTCTAGAAATTTGGTTTTATGAAATGCCCCCTCCCTGATTTGGCCACAG	20577
Db	1693	CATAAAACAGCTCTAGAAATTTGGTTTTATGAAATGCCCCCTCCCTGATTTGGCCACAG	17522
Qy	2058	CATGATGACATGTTGGCCCTTGTTTCTAGTAAGAACAATCGTGCCCTCCCTCCCTGAGA	21177
Db	1753	CATGATGACATGTTGGCCCTTGTTTCTAGTAAGAACAATCGTGCCCTCCCTCCCTGAGA	18122
Qy	2118	ACTGTTAAGTTCTTATTTTAAGCTCTTCTCGAGACTAATGAACATGATGAGAGCCTATTAATA	21777
Db	1813	ACTGTTAAGTTCTTATTTTAAGCTCTTCTCGAGACTAATGAACATGATGAGAGCCTATTAATA	18722
Qy	2178	TGTCCCAACAGTTTCAATTTTGGCCATTTGGAACCAATCATTTGATTTTAAAGTGAATT	22377
Db	1873	TGTCCCAACAGTTTCAATTTTGGCCATTTGGAACCAATCATTTGATTTTAAAGTGAATT	19322
Qy	2238	ATCTTT--GAAAAACATTTATTTATTTACCTTACCTACAGA	2269
Db	1933	ATCTTTGTTTTCATTTATTTATTTACCGGAAAGA	1966
RESULT 3			
US-10-183-116-15			
Sequence 15, Application US/10183116			
Publication No. US20030092035A1			
GENERAL INFORMATION:			
APPLICANT: Anderson, David J.			
APPLICANT: Dong, Xinzhong			
APPLICANT: Zylka, Mark			
APPLICANT: Simon, Melvin			
APPLICANT: Han, Sang-kyou			
TITLE OF INVENTION: PAIN SIGNALING MOLECULES			
FILE REFERENCE: CALTE, AICIPI			
CURRENT APPLICATION NUMBER: US/10/183,116			
CURRENT FILING DATE: 2002-06-26			
PRIOR APPLICATION NUMBER: US 60/222,344			
PRIOR FILING DATE: 2000-08-01			
PRIOR APPLICATION NUMBER: US 60/202,027			
PRIOR FILING DATE: 2000-05-04			
PRIOR APPLICATION NUMBER: US 09/704,707			
PRIOR FILING DATE: 2000-11-03			
PRIOR APPLICATION NUMBER: US 60/285,493			
PRIOR FILING DATE: 2001-04-19			
PRIOR APPLICATION NUMBER: US 09/849,869			
PRIOR FILING DATE: 2001-05-04			
NUMBER OF SEQ ID NOS: 109			
SOFTWARE: FastSeq for Windows Version 4.0			
SEQ ID NO 15			
LENGTH: 2040			
TYPE: DNA			
ORGANISM: Homo sapiens			
FEATURE:			
NAME/KEY: CDS			
LOCATION: (328) ... (1293)			
US-10-183-116-15			

Query Match	56.7%;	Score 1483.8;	DB 14;	Length 2040;
Best Local Similarity	91.4%;	Pred. No. 0;		
Matches 1596;	Conservative	0;	Mismatches 147;	Indels 4;
				Gaps 2;

QY	4432	TCATGGGATCAAGACTGGGGTTTCTGAGATGATTAACCAATCCAGTCTTGAGTCA	521
Db	298	TCAGAGGTCACCAAGACTAGGGTTTCTAGCATGATTCACCACTCTCAACTTGGACACA	357
QY	552	GAATGACCAACCAATCAACGACGTGAGGAGACTCTCTTGCTACAAAGCAGACCTTGAGCTTC	581
Db	358	GAATGACCAACCAATCAACGAGAACTGAGGAGACTCTTGCTACAAAGACCACTTGAGCTTC	417
QY	582	ACGGGGCTGAGGTGATCGTTTCCCTGTCGGCGGTCAGAGAAACGGGGTGTGTGCG	641
Db	418	ACGGGTGAGGTGATCGTTTCCCTGTCGGCGGTCAGAGAAACGAGTTGTGTCTTG	477
QY	642	CTCTGGGCTGCCGATGCGGAGAGACGCTGTCTTCATCTACATCTCAACCTGGTCGCG	701
Db	478	CTCTGGGCTGCCGATGCGGAGAGCGCTTCTTCATCTACATCTCAACTTGGCGCGCA	537
QY	702	GCCGACTTCTCTTCTCTTAAGGGCCACTTATATGTTGCGCGTTAAGCTCATCATATTC	761
Db	538	GCAAGCTTCTCTTCTCTCAGGGCGCGCTTATATATTCCTGTTAAGCTTATCATAGATTC	597
QY	762	CGCCATCCCATCTCCAAATCCTCAGTCTGTGATGACCTTCTCCCTATTTATAGCCCTA	821
Db	598	CCCATACCATCTCTAAATCTCTATCTGTGATGATGTTTCTTAATTTGCAAGCCCTG	657
QY	822	AGCATGTGACGCGCCATCAGACCGAGCGGTGCTGTTCATCTGTGGCCATCTGTATC	881
Db	658	AGCTTCTGATGCGCGTATGACACGAGCGCGTCTGTCTGTGTGCGCCATCTGTATC	717
QY	882	CACTCGCGCGCGCCAGATACCTGTCAATCGGTCAATGTGTGTCTGTGAGCCCTGTGCG	941
Db	718	CGCTGCAACCGGCCACACACCTGTCAACGGGTGTGTGTGTCTGTGCTTGGAGCCCTGTGCG	777
QY	942	CTGCTGCGAGATCTCTGAGTGTATGTTCTGTGACTTCTCTGTTATGATGTGATTCCT	1001
Db	778	CTGCTGCGAGATCTCTGAGTGTATGTTATGTGACTTCTCTGTTACATGTGTGATTCCT	837
QY	1002	GTTTGTGTGAAAGCTCAGATTTTCAATTCAATCCGCTGTGGCTGTTTTTATATGTGTGT	1061
Db	838	GCTTGTGTGAAACATCAGATTTTCAACATCGGTGTGTGATTTTATATGTGTGT	897
QY	1062	CTGTGTGGGTCAAGCTGATCTGTGCTGTGTCAGAGATTCCTGTGATCCCGAAGATGTGCG	1121
Db	898	CTGTGTGGGTCAAGCTGATCTGTGCTGTGTCAGAGATTCCTGTGATCCCGAAGATGTGCG	957
QY	1122	CTGACAGAGCTGTACGTGACCACTCTCTCAAGTGTGTTCTTCTCTGTGTGGCTG	1181
Db	958	CTGACAGAGCTGTACGTGACCACTCTCTCTCAAGTGTGTTCTTCTCTGTGTGGCTG	1017
QY	1182	CCCTTGGCATTCAGTGGGCGCTGTTTTCCAGATCCACCTGTGATTTGGAAGTCTTATTT	1241
Db	1018	CCCTTGGCATTCAGTGTGTTTTCTTATTTTATGAGATCCAGTGAACGGAAGTCTTATTT	1077
QY	1242	TGTCAATGTGACATCTAGTTTCAATTTTCTGTCCGCTCTTAAACAGAGTGCCAACCCATC	1301
Db	1078	TGTCAATGTGACATCTAGTTTCTAATTTTCTGTCCGCTCTTAAACAGAGTGCCAACCCATC	1137
QY	1302	ATTATCTTCTGTGTGGGCTCTCTTATAGGACGGGTCAAAATATAGCAGAACTGTGATGTGTT	1361
Db	1138	ATTATCTTCTGTGTGGGCTCTCTTATAGGACGGGTCAAAATATAGCAGAACTGTGATGTGTT	1197
QY	1362	CTCCAGAGGGCTCTGACAGACACGCTGAGGTGATGAAAGGTGAGGGGTGTGCTTCTCAG	1421
Db	1198	CTCCAGAGGGCTCTGACAGAGCGCTCTGAGGTGATGAAAGGTGAGGGGACCTTCTGTAG	1257
QY	1422	GAAGCCCTGAGCTGTGCGGAAGCAGATTGAGCAGTGAAGAGAACTCTGTGCTGTGCA	1481
Db	1258	GAAATCTGTAGAGCTGTGCGGAAGCAGATTGAGCAGTGAAGAGAGCTCTGTGCTGTGCA	1317
QY	1482	GACAGGACTTTGAGAGCAATGTGTGCTGTGCCACCTGTGACATTTATATGATTTTTCTTA	1541
Db	1318	GACAGGACTTTGAGAGCAACTGTGCTGTGCCACCTGTGACATTTATATGAGTTTTCTTA	1377

QY	1542	GCCTTCGCCCTCAGAAAAGTCTCAGTGGTCCCTCAAGGCTCTTGGAAATGATGTTATCTA	1601
Db	1378	GCTTCTGCTCTCAGAAATGTCCTCAGTGGTCTCTCAAGGCTCTTCAATATGATGTTATCTA	1437
QY	1602	ACCTGACAGTTGACAGTTTTCACCCATGGAAAGCATTAAGTCTGACAGTACCAATGTTGGAT	1661
Db	1438	ACCTGACAGTTGGCGGTTTTCACCCATGGAAAGCATTAAGTCTGACAGTACCAATGTTGGAT	1497
QY	1662	TCTTCCTTGATATTAACCAATACATTTTCCCTGTATCTTGACCTGATTTTTCCTACTGA	1721
Db	1498	TCTTCCTTGATATTAACCAACATATTTCCCTGTATCTTGACCTGATTTTTCCTACTGA	1557
QY	1722	CACATTTTTCGACATTTTTCATTTGTAATTAAGAGGTTGCTGTCCACAAACCTTAACCTCT	1781
Db	1558	CACATTTTTCGCAATTTTCTTTGTAATTAAGAGGTTGCTGTACAAACCTTAACCAAC-	1616
QY	1782	TCTTTATATCTGTTTTCCTCACTGATAGTATCAAAAAGGAAGATTCTTAATATCTGTCA	1841
Db	1617	TCTTTATATCTTCTTCTTCACTGATAGTATCAAAAAGGAAGATTCTTAATATCTGTCA	1676
QY	1842	GACTATGTTCCCTCGAAAATCATGTTCCCTTTATGACTGAGGCATTACTGCAAGTTGGA	1901
Db	1677	GACTATGTTCCCTCGAAAATCATGTTCCCTTTATGACTGAGGCATTACTGCAAGTTGGA	1736
QY	1902	AGCTCAATTTCTTAATAGTGAAGTTCTGTACTCTTAATTCATTTGAATTCCTCAGATTA	1961
Db	1737	AGCTCGATTTCTTAATAGTGAAGTTCTGTACTCTCTACATTCATTTGAATTCCTCAGATTA	1796
QY	1962	AMGCAAAATATAGACCTTAGAGAGAGATTTCCTCTCATTAATAAACAGCTTAGAAATGG	2021
Db	1797	GAGCAAAATATATGTCCTTAGAGAGAGACTCTCTCTTCATTAATAAACAGCTTAGAATGG	1856
QY	2022	TTTTATGAAATAGCCCTCTCTGTCAATTTGTCACAGCAATGTGAATGTTGGCTTGTT	2081
Db	1857	TTTTATGAAATAGCTTCCCTGTCAATTTGTTCAAGCAATGTGAATGTTGGCTTGTT	1916
QY	2082	TCTAGTAAAGCAATCTGTGGGCCCTTCCCTCTTGAGAACTGGTAAGTTCTTAATTAAGCTCT	2141
Db	1917	TCTAGTAAAGCAACTGTGGGCCCTTCCCTCTTGAGAACTTTTAAGTGTCTTAATTAAGCTCT	1976
QY	2142	TCCGGACTAATGAACTAGTAGGAGGAGCTAATAAATATGTCACACAGTTTCATTTTGACC	2201
Db	1977	TCCGGACTAATGAACTAGTAGGAGGAGGAGCTAATAAATGTC---CCCAAGTTCTAATTTTGACC	2033
QY	2202	ATTGGA 2208	
Db	2034	ATTGGA 2040	
RESULT 4			
US-10-292-798-1273			
; Sequence 1273, Application US/10292798			
; Publication No. US20030235833A1			
GENERAL INFORMATION:			
; APPLICANT: SUMA, MAKIKO			
; APPLICANT: ASAI, KIYOSHI			
; APPLICANT: AKIYAMA, YUTAKA			
; APPLICANT: ABRAMANI, HIROYUKI			
; TITLE OF INVENTION: GUANOSINE TRIPHOSPHATE-BINDING PROTEIN COUPLED RECEPTORS			
; FILE REFERENCE: 084335/166			
; CURRENT APPLICATION NUMBER: US/10/392, 798			
; CURRENT FILING DATE: 2002-11-13			
; PRIOR APPLICATION NUMBER: 10/017,161			
; PRIOR FILING DATE: 2001-12-18			
; PRIOR APPLICATION NUMBER: JP 2001-246789			
; PRIOR FILING DATE: 2001-06-18			
; NUMBER OF SEQ ID NOS: 2070			
; SOFTWARE: PatentIn Ver. 2.1			
; SEQ ID NO 1273			
; LENGTH: 1369			
; TYPE: DNA			
; ORGANISM: Homo sapiens			
; FEATURE:			

[illegible]

	1071	1422	1131	1482	1191	1542	1251	1602	1311
Db	CTCCAGAGGGCTCTGCAAGACACGGCTGAGTGTGATGAAGGTGAGGGTGGTTCTCTCAG								
Qy	GAAACCCCTGAGACTGTCCGGGAAGCGAATTGGAGCGATGAGGAAAGAACTCTGCCCTGTCA	1481							
Db	GAAACCCCTGAGACTGTCCGGGAAGCGAATTGGAGCGATGAGGAAAGAACTCTGCCCTGTCA	1190							
Qy	GACAGGACTTTGAGAGCAATGCTGCGCCCTGCCACCCCTTGACAAATTATATGACTTTTCTTA	1541							
Db	GACAGGACTTTGAGAGCAATGCTGCGCCCTGCCACCCCTTGACAAATTATATGACTTTTCTTA	1250							
Qy	GCTTCTTCCTCAGAAATGTCTCAGATGTGCTCCCTCAAGGCTTCGAAATGATGATTAATCTA	1601							
Db	GCTTCTTCCTCAGAAATGTCTCAGATGTGCTCCCTCAAGGCTTCGAAATGATGATTAATCTA	1310							
Qy	ACCTGACAGTTGCAATTTTTCACCCATGGAAGCAATTAATCTGACACGTACAAATGTTTGGCA	1660							
Db	ACCTGACAGTTGCAATTTTTCACCCATGGAAGCAATTAATCTGACACGTACAAATGTTTGGCA	1369							


```

Db      662 AGCATGCTGAGGCCATCAGACCGAGCGCTGCTGTCATCTGTGGCCATCTGTGAC 721
Qy      882 CACTGCCGCGCCCCAGATACCTGTCACTGTCATGTGTGTCCTGTGGCCCTGTGC 941
Db      722 CACTGCCGCGCCCCAGATACCTGTCACTGTCATGTGTGTCCTGTGGCCCTGTGC 781
Qy      942 CTGCTGCGGAGTATCTGTGAGTGAATGTTCTGTGACCTTCTGTGTTAGTGTGCTGATTCT 1001
Db      782 CTGCTGCGGAGTATCTGTGAGTGAATGTTCTGTGACCTTCTGTGTTAGTGTGCTGATTCT 841
Qy      1002 GTTTGGTGAAGAGTCAGATTTTCAATTCATTCGAGTGGCTGTTTATAGTGTGTT 1061
Db      842 GTTTGGTGAAGAGTCAGATTTTCAATTCATTCGAGTGGCTGTTTATAGTGTGTT 901
Qy      1062 CTCTGTGGTTCACGCTGTGTCCTGTGTCAGATTTCTGTGATTCGCGAAGATGCG 1121
Db      902 CTCTGTGGTTCACGCTGTGTCCTGTGTCAGATTTCTGTGATTCGCGAAGATGCG 961
Qy      1122 CTGACAGGCTGTACGTGACCAATCTCTCAAGTGTGTTCTCTCTGTGGCTG 1181
Db      962 CTGACAGGCTGTACGTGACCAATCTCTCAAGTGTGTTCTCTCTGTGGCTG 1021
Qy      1182 CCGTTTGGCAATCTAGTGGGCGCTGTTTCCAGATCCACCTGGATTTGGAAGCTTATTT 1241
Db      1022 CCGTTTGGCAATCTAGTGGGCGCTGTTTCCAGATCCACCTGGATTTGGAAGCTTATTT 1081
Qy      1242 TGTCAATGTGATCTAGTTTCCATTTTCTGTCCGCTTTAAACAGAGTCCACCCATC 1301
Db      1082 TGTCAATGTGATCTAGTTTCCATTTTCTGTCCGCTTTAAACAGAGTCCACCCATC 1141
Qy      1302 ATTACTTCTTCTGTGGGCTCTTTAGGCAAGCTCAAAATAGGCAAGCTGAGTGT 1361
Db      1142 ATTACTTCTTCTGTGGGCTCTTTAGGCAAGCTCAAAATAGGCAAGCTGAGTGT 1201
Qy      1362 CTCAGAGGGCTCTGAGGACGACGCTGAGGTGATGTAAGGTGAGGGGTCTCTGAG 1421
Db      1202 CTCAGAGGGCTCTGAGGACGACGCTGAGGTGATGTAAGGTGAGGGGTCTCTGAG 1261
Qy      1422 GAAACCTGTGAGCTGTGCGGAAGCAGATTGAGAGCAGTGAAGAACTGTGCTGTCA 1481
Db      1262 GAAACCTGTGAGCTGTGCGGAAGCAGATTGAGAGCAGTGAAGAACTGTGCTGTCA 1321
Qy      1482 GACAGAGCTTTGAGAGCAATGCTGCTGCCACCTTGACAATTATATGCAATTTTCTTA 1541
Db      1322 GACAGAGCTTTGAGAGCAATGCTGCTGCCACCTTGACAATTATATGCAATTTTCTTA 1381
Qy      1542 GCCTTCGCTCAGAAATG 1560
Db      1382 GCCTTCGCTCAGAAATG 1400

```

RESULT 8

```

US-10-017-161-1599
; Sequence 1599, Application US/10017161
; Publication No. US20030143688A1
; GENERAL INFORMATION:
; APPLICANT: SUMA, MAKIKO
; APPLICANT: ASAI, KIYOSHI
; APPLICANT: AKIYAMA, YUTAKA
; APPLICANT: AUBURANT, HIROYUKI
; TITLE OF INVENTION: NOVEL G PROTEIN-COUPLED RECEPTORS
; FILE REFERENCE: 084335/0152
; CURRENT APPLICATION NUMBER: US/10/017.161
; CURRENT FILING DATE: 2002-12-18
; PRIOR APPLICATION NUMBER: JP 2001/246789
; PRIOR FILING DATE: 2001-06-18
; NUMBER OF SEQ ID NOS: 2430
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 1599
; LENGTH: 1370
; TYPE: DNA
; ORGANISM: Homo sapiens
; FEATURE:

```

```

; NAME/KEY: source
; LOCATION: (1) .. (1370)
; FEATURE:
; NAME/KEY: CDS
; LOCATION: (201) .. (619)
; FEATURE:
; NAME/KEY: CDS
; LOCATION: (966) .. (1170)
; US-10-017-161-1599

```

```

Query Match 40.2%; Score 1053; DB 15; Length 1370;
Best Local Similarity 94.4%; Pred. No. 1.5e-286;
Matches 1136; Conservative 0; Mismatches 60; Indels 7; Gaps 4;

```

```

Qy      462 TCATGGGTCAATCAGATGGGGTTCATGAGATGATTCACATCCAGTCTGGGTACA 521
Db      171 TCAGGGTCAATCAGATGGGGTTCATGAGATGATTCACATCCAGTCTGGGTACA 230
Qy      522 GAATGACCAATCAACGAGCGTGAAGAGATCTCTTGTCTAACAGCAGACCTGAGCTTC 581
Db      231 GAATGACCAATCAACGAGCGTGAAGAGATCTCTTGTCTAACAGCAGACCTGAGCTTC 290
Qy      582 ACGGGGCTGAGTGAATGTTTCCCTTGTGCGCGCTGACAGGAAACGGGTTGTGCTG 641
Db      291 ACGGGGCTGAGTGAATGTTTCCCTTGTGCGCGCTGACAGGAAACGGGTTGTGCTG 350
Qy      642 CTCTGGGCTGCGCATGCGAGAGACGTGTCTCATCTACATCTCTCAACTGTGCG 701
Db      351 CTCTGGGCTGCGCATGCGAGAGACGTGTCTCATCTACATCTCTCAACTGTGCG 410
Qy      702 GCCGACTTCTCTTCTTAAAGGCGCAATTATATGTTGCGGTTAGCCCTCATCAATATC 761
Db      411 GCCGACTTCTCTTCTTAAAGGCGCAATTATATGTTGCGGTTAGCCCTCATCAATATC 470
Qy      762 GCGCATCCCATCTCAAAATCTCAGTCTGTGATGACCTTTCCCTTATAGGCTTA 821
Db      471 GCGCATCCCATCTCAAAATCTCAGTCTGTGATGACCTTTCCCTTATAGGCTTA 530
Qy      822 AGCATGCTGAGCGCATCAGACCGAGCGCTGCTGTCAATCTGTGGCCATCTGTGAC 881
Db      531 AGCATGCTGAGCGCATCAGACCGAGCGCTGCTGTCAATCTGTGGCCATCTGTGAC 590
Qy      882 CACTGCCGCGCCCCAGATACCTGTCAATCGTCAATGTGTCTGTGCTGTGGCCCTGTGC 941
Db      591 CACTGCCGCGCCCCAGATACCTGTCAATCGTCAATGTGTCTGTGCTGTGGCCCTGTGC 649
Qy      942 CTGCGGCGGAGTATCTGAGTGAATGTTCTGTGATCTCTGTTTATGAGTGTGATTCT 1001
Db      650 CTGCGGCGGAGTATCTGAGTGAATGTTCTGTGATCTCTGTTTATGAGTGTGATTCT 709
Qy      1002 GTTTGGTGAAGAGTCAGATTTTCAATTCATTCGAGTGGCTGTTTATAGTGTGTT 1061
Db      710 GTTGGTGAAGAGTCAGATTTTCAATTCATTCGAGTGGCTGTTTATAGTGTGTT 769
Qy      1062 CT---CTGTGGTTCAGCTGTGTCTGTGTGATGAGATTTCTGTGATTCGCGAAGATG 1118
Db      770 CTGTGCGGGGGGTCAAGGTGTCTGTCTCATGAGATGAGCGGATATGCGGAAATG 829
Qy      1119 CGGCTGACAGGCT-GTACGAGCAATCTCTCAGAGTGTGCTGTCTCTCTCTGTG 1177
Db      830 CGGCTGAGATGTTCAAGCGTATCTCTCTCATATGTGTGTCTCTCTCTGCGGTGT 889
Qy      1178 CCGTCCCTTTGGCAATCAGTGGGCTGTGTTTCCAGATCCACCTGATTTGGAAGTCTT 1237
Db      890 CCGTCCCTTTGGCAATCAGTGGGCTGTGTTTCCAGATCCACCTGATTTGGAAGTCTT 947
Qy      1238 ATTTGTGATGTGATCTAGTTTCAATTTTCTGTGCTGTCTTTAAACAGCAGTGTCAACC 1297
Db      948 ATTTGTGATGTGATCTAGTTTCAATTTTCTGTGCTGTCTTTAAACAGCAGTGTCAACC 1007
Qy      1298 CATGATTTACTTCTTGTGGGCTCTTTAGGCAAGCTCAAAATAGGAGAACGTAAGCT 1357
Db      1008 CATGATTTACTTCTTGTGGGCTCTTTAGGCAAGCTCAAAATAGGAGAACGTAAGCT 1067

```


Query Match 38.7%; Score 1013.6; DB 14; Length 1604;
 Best Local Similarity 91.0%; Pred. No. 2,4e-275;
 Matches 1090; Conservative 0; Mismatches 104; Indels 4; Gaps 1;

QY 466 GGGTCATCAGATGGGGTTTCTGAGCATGAGATTCAACCATCCAGTCTGGGTACAGAAC 525
 DB 407 GGGGCACACAGACTAGAGTTTCTGAGCATGAGATCCACCGTCCAGTCTTGGTACAAAC 466
 QY 526 TGAACCAATCAACGAGCTGAGAGACTCTCTTCTCAAGCAGACCCCTGAGCTTACCG 585
 DB 467 TGACCAATCAACGAGCTGAGAGACTCTCTTCTCAAGCAGACCCCTGAGCTTACCG 526
 QY 586 GGCCTGAGTCAATGCTTCCCTTGTGCGGCTGACAGGAAACGGGTTGTCTGAGCTCC 645
 DB 527 TGTCTGAGTCAATGCTTCCCTTGTGCGGCTGACAGGAAACGGGTTGTCTGAGCTCC 586
 QY 646 TGGGCTGCGGCTGAGAGAGAGCTGTCTCAATCAATCAATCAATCAATCAATCAATCA 705
 DB 587 TGGGCTGAGCTGAGAGAGAGCTGTCTCAATCAATCAATCAATCAATCAATCAATCA 646
 QY 706 ACTTCTCTCTTCAAGGCGGCAATTAATGTCGCGGTACCGCTCAATTAATCCGCC 765
 DB 647 ACTTCTCTCTTCAAGTCTTCAATTAATGTCGCGGTACCGCTCAATTAATCCGCC 706
 QY 766 ATCCCATCTCCAAATCTCAGTCTGATGATGATGATGATGATGATGATGATGATGATG 825
 DB 707 ATCTCATCCGCAAAATCTCTGTTTCTGATGATGATGATGATGATGATGATGATGATG 766
 QY 826 TGTCTGAGCGGCTGAGAGAGAGCTGTCTCAATCAATCAATCAATCAATCAATCAATCA 885
 DB 767 TGTCTGAGCGGCTGAGAGAGAGCTGTCTCAATCAATCAATCAATCAATCAATCAATCA 826
 QY 886 GCGCGCGCGGCAATGATGATGATGATGATGATGATGATGATGATGATGATGATGATG 945
 DB 827 GCGCGCGCGGCAATGATGATGATGATGATGATGATGATGATGATGATGATGATGATG 886
 QY 946 TGGGAGTATCTGAGTGTGATGATGATGATGATGATGATGATGATGATGATGATGATG 1005
 DB 887 TGTGTGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATG 946
 QY 1006 GGTGTGAAACGTGATGATGATGATGATGATGATGATGATGATGATGATGATGATG 1065
 DB 947 GGTGTGAAACGTGATGATGATGATGATGATGATGATGATGATGATGATGATGATG 1006
 QY 1066 GTGGGTCACAGCTGTGTCTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGT 1125
 DB 1007 GT 1066
 QY 1126 CGAGGCTGTGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATG 1185
 DB 1067 CGAGGCTGTGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATG 1126
 QY 1186 TTGGGATTCAGTGTGATGATGATGATGATGATGATGATGATGATGATGATGATGATG 1245
 DB 1127 TTGGGATTCAGTGTGATGATGATGATGATGATGATGATGATGATGATGATGATGATG 1186
 QY 1246 ATGTGTGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 1305
 DB 1187 ATGTGTGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 1246
 QY 1306 ACTTCTCTCTTCAAGGCGGCAATTAATGTCGCGGTACCGCTCAATTAATCCGCC 1365
 DB 1247 ACTTCTCTCTTCAAGGCGGCAATTAATGTCGCGGTACCGCTCAATTAATCCGCC 1306
 QY 1366 AGAGGGCTGTGAGAGAGAGCTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGT 1425
 DB 1307 AGAGGGCTGTGAGAGAGAGCTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGT 1366
 QY 1426 CCTGTGAGCTGT 1481
 DB 1367 GCGTGTGAGCTGT 1426
 QY 1482 GACAGGATCTTGAAGAGATGATGATGATGATGATGATGATGATGATGATGATGATG 1541

Db 1427 GACGGGATCTTGAAGAGATGATGATGATGATGATGATGATGATGATGATGATGATG 1486
 QY 1542 GCGTGTGAGCTGT 1601
 DB 1487 GCGTGTGAGCTGT 1546
 QY 1602 ACCTGACATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATG 1659
 DB 1547 ACCTGACATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATG 1604

RESULT 12
 US-10-225-567A-688
 ; Sequence 688, Application US/10225567A
 ; Publication No. US20030113798A1
 ; GENERAL INFORMATION:
 ; APPLICANT: LifeSpan Biosciences
 ; APPLICANT: Brown, Joseph P.
 ; APPLICANT: Burner, Glenna C.
 ; APPLICANT: Roush, Christine L.
 ; TITLE OF INVENTION: ANTIGENIC PEPTIDES AND ANTIBODIES FOR G PROTEIN-COUPLED RECEPTORS
 ; FILE REFERENCE: 1920-4-4
 ; CURRENT APPLICATION NUMBER: US/10/225,567A
 ; CURRENT FILING DATE: 2001-12-19
 ; PRIOR APPLICATION NUMBER: 60/257,144
 ; PRIOR FILING DATE: 2000-12-19
 ; NUMBER OF SEQ ID NOS: 2292
 ; SOFTWARE: PatentIn version 3.1
 ; SEQ ID NO 688
 ; LENGTH: 1604
 ; TYPE: DNA
 ; ORGANISM: Homo sapiens
 US-10-225-567A-688

Query Match 38.7%; Score 1013.6; DB 15; Length 1604;
 Best Local Similarity 91.0%; Pred. No. 2,4e-275;
 Matches 1090; Conservative 0; Mismatches 104; Indels 4; Gaps 1;

QY 466 GGGTCATCAGATGGGGTTTCTGAGCATGAGATTCAACCATCCAGTCTGGGTACAGAAC 525
 DB 407 GGGGCACACAGACTAGAGTTTCTGAGCATGAGATCCACCGTCCAGTCTTGGTACAAAC 466
 QY 526 TGAACCAATCAACGAGAGAGAGCTGTCTCAATCAATCAATCAATCAATCAATCAATCA 585
 DB 467 TGACCAATCAACGAGAGAGAGCTGTCTCAATCAATCAATCAATCAATCAATCAATCA 526
 QY 586 GGCCTGAGTCAATGCTTCCCTTGTGCGGCTGACAGGAAACGGGTTGTCTGAGCTCC 645
 DB 527 TGTCTGAGTCAATGCTTCCCTTGTGCGGCTGACAGGAAACGGGTTGTCTGAGCTCC 586
 QY 646 TGGGCTGCGGCTGAGAGAGAGCTGTCTCAATCAATCAATCAATCAATCAATCAATCA 705
 DB 587 TGGGCTGAGCTGAGAGAGAGCTGTCTCAATCAATCAATCAATCAATCAATCAATCA 646
 QY 706 ACTTCTCTCTTCAAGGCGGCAATTAATGTCGCGGTACCGCTCAATTAATCCGCC 765
 DB 647 ACTTCTCTCTTCAAGTCTTCAATTAATGTCGCGGTACCGCTCAATTAATCCGCC 706
 QY 766 ATCCCATCTCCAAATCTCAGTCTGATGATGATGATGATGATGATGATGATGATGATG 825
 DB 707 ATCTCATCCGCAAAATCTCTGTTTCTGATGATGATGATGATGATGATGATGATGATG 766
 QY 826 TGTCTGAGCGGCTGAGAGAGAGCTGTCTCAATCAATCAATCAATCAATCAATCAATCA 885
 DB 767 TGTCTGAGCGGCTGAGAGAGAGCTGTCTCAATCAATCAATCAATCAATCAATCAATCA 826
 QY 886 GCGCGCGCGGCAATGATGATGATGATGATGATGATGATGATGATGATGATGATGATG 945
 DB 827 GCGCGCGCGGCAATGATGATGATGATGATGATGATGATGATGATGATGATGATGATG 886
 QY 946 TGGGAGTATCTGAGTGTGATGATGATGATGATGATGATGATGATGATGATGATGATG 1005

APPLICANT: Chu, Zhi Liang
 APPLICANT: Dang, Huang T.
 APPLICANT: Lowitz, Kevin P.
 APPLICANT: Priddy, Cameron
 TITLE OF INVENTION: Endogenous And No. US2003013588A9-Endogenous Versions of Human C
 TITLE OF INVENTION: Receptors
 FILE REFERENCE: AREN-0308
 CURRENT APPLICATION NUMBER: US/09/995,225
 CURRENT FILING DATE: 2001-11-26
 PRIOR APPLICATION NUMBER: 09/170,496
 PRIOR FILING DATE: 1998-10-13
 PRIOR APPLICATION NUMBER: PCT/US99/23938
 PRIOR FILING DATE: 1998-10-13
 PRIOR APPLICATION NUMBER: 60/253,404
 PRIOR FILING DATE: 2000-11-27
 PRIOR APPLICATION NUMBER: 60/255,366
 PRIOR FILING DATE: 2000-12-12
 PRIOR APPLICATION NUMBER: 60/270,286
 PRIOR FILING DATE: 2001-02-20
 PRIOR APPLICATION NUMBER: 60/282,365
 PRIOR FILING DATE: 2001-04-06
 PRIOR APPLICATION NUMBER: 60/270,266
 PRIOR FILING DATE: 2001-02-20
 PRIOR APPLICATION NUMBER: 60/282,032
 PRIOR FILING DATE: 2001-04-06
 PRIOR APPLICATION NUMBER: 60/282,358
 PRIOR FILING DATE: 2001-04-06
 PRIOR APPLICATION NUMBER: 60/282,356
 PRIOR FILING DATE: 2001-04-06
 PRIOR APPLICATION NUMBER: 60/290,917
 PRIOR FILING DATE: 2001-05-14
 PRIOR APPLICATION NUMBER: 60/309,208
 PRIOR FILING DATE: 2001-07-31
 NUMBER OF SEQ ID NOS: 67
 SOFTWARE: PatentIn version 3.1
 SEQ ID NO 19
 LENGTH: 969
 TYPE: DNA
 ORGANISM: Artificial Sequence
 FEATURE:
 OTHER INFORMATION: No. US2003013588A9el Sequence
 US-09-995-225-19

Query Match 37.0%; Score 969; DB 10; Length 969;
 Best Local Similarity 100.0%; Pred. No. 7.2e-263;
 Matches 969; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

492 ATGATTCACATCCAGCTGTTGGGTACAGAACTGACCAATCAACGACGTGAGAG 551
 1 ATGATTCACATCCAGCTGTTGGGTACAGAACTGACCAATCAACGACGTGAGAG 60
 552 ACTCCCTGCTACAGCAAGCACTGAGCTTCAAGGGCTGAGTGCATGCTTCCCTGTC 611
 61 ACTCCCTGCTACAGCAAGCACTGAGCTTCAAGGGCTGAGTGCATGCTTCCCTGTC 120
 612 GCGCTGACAGAAAGCGGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 671
 121 GCGCTGACAGAAAGCGGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 180
 672 GCTCCATCTACATCTCAACCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 731
 181 GCTCCATCTACATCTCAACCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 240
 732 ATATGTTCCGCTGCTACAGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 791
 241 ATATGTTCCGCTGCTACAGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 300
 792 GTGATGACCTTCCCTACTTATAGGCTTAAGCATGCTGAGGCGCATGACGACGAGGCG 851
 301 GTGATGACCTTCCCTACTTATAGGCTTAAGCATGCTGAGGCGCATGACGACGAGGCG 360
 852 TGCCGTGCATCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 911

361 TGCCGTGCATCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 420
 912 GTGATGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 971
 421 GTGATGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 480
 972 TGATGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 1031
 481 TGATGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 540
 1032 ATCCGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 1091
 541 ATCCGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 600
 1092 AGATTCCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 1151
 601 AGATTCCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 660
 1152 AGATTCCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 1211
 661 AGATTCCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 720
 1212 AGATTCCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 1271
 721 AGATTCCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 780
 1272 TCCGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 1331
 781 TCCGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 840
 1332 CGTCAAAATAGGCAAGAACTGAGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 1391
 841 CGTCAAAATAGGCAAGAACTGAGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 900
 1392 GTGATGAGGTGAGGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 1451
 901 GTGATGAGGTGAGGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 960
 1452 GAGCAGTGA 1460
 961 GAGCAGTGA 969

Search completed: October 29, 2004, 19:12:59
 Job time : 1223 secs

GenCore version 5.1.6
Copyright (c) 1993 - 2004 CompuGen Ltd.

OM nucleic - nucleic search, using sw model

Run on: October 29, 2004, 09:26:47 ; Search time 10822 Seconds
(without alignments)
1140.070 Million cell updates/sec

Title: US-09-867-570-1

Perfect score: 2618
Sequence: 1 aacacatgcgcgcgaattcgcg.....gaaaaaaaaaaaaaaaaa 2618

Scoring table: IDENTITY_NUC
Gapop 10.0 , Gapext 1.0

Searched: 4526729 seqs, 23644849745 residues

Total number of hits satisfying chosen parameters: 9053458

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%
Listing first 45 summaries

Database :

GenEmbl:.*
1: gb_ba:.*
2: gb_ptg:.*
3: gb_in:.*
4: gb_cm:.*
5: gb_ov:.*
6: gb_pat:.*
7: gb_ph:.*
8: gb_dl:.*
9: gb_pr:.*
10: gb_ro:.*
11: gb_sts:.*
12: gb_sy:.*
13: gb_un:.*
14: gb_vi:.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB	ID	Description
1	2618	100.0	2618	6	AX429465	AX429465 Sequence
2	2128.6	81.3	91923	9	AC108007	AC108007 Homo sapi
3	2128.6	81.3	172939	9	AC090099	AC090099 Homo sapi
4	1987.4	75.9	192618	2	AC020568	AC020568 Homo sapi
5	1842.8	70.4	156839	9	AC107948	AC107948 Homo sapi
6	1792.2	68.5	163718	9	AC023078	AC023078 Homo sapi
7	1792.2	68.5	165490	2	AC139482	AC139482 Homo sapi
8	1637.2	62.5	165490	2	AC139482	AC139482 Homo sapi
9	1483.8	56.7	2040	6	CO815414	CO815414 Sequence
10	1483.8	56.7	2040	6	AX299143	AX299143 Sequence
11	1483.8	56.7	2040	9	AY042213	AY042213 Homo sapi
12	1300.2	49.7	1697	9	BC067292	BC067292 Homo sapi
13	1228.2	46.9	189230	9	AC103974	AC103974 Homo sapi
14	1192.6	45.6	1369	6	AX647081	AX647081 Sequence
15	1104	42.2	8622	6	AX429467	AX429467 Sequence
16	1094.2	41.8	1400	6	AX299158	AX299158 Sequence
17	1094.2	41.8	1400	6	AX549388	AX549388 Sequence
18	1094.2	41.8	1400	9	AY042215	AY042215 Homo sapi
19	1040.6	39.7	1369	6	AX646705	AX646705 Sequence

20	1040.6	39.7	1369	9	AB065846	AB065846 Homo sapi
21	1013.6	38.7	1604	6	AX299160	AX299160 Sequence
22	1013.6	38.7	1604	6	AX549403	AX549403 Sequence
23	1013.6	38.7	1604	9	AY042216	AY042216 Homo sapi
24	1012.2	38.7	1369	6	AX646849	AX646849 Sequence
25	969	37.0	969	6	E43450	E43450 Novel prote
26	969	37.0	969	6	AX498196	AX498196 Sequence
27	967.4	37.0	969	6	E43451	E43451 Novel prote
28	965.8	36.9	969	6	AX923125	AX923125 Sequence
29	965.6	36.9	1116	6	BD095749	BD095749 Novel gua
30	961	36.7	969	9	AF474987	AF474987 Homo sapi
31	949.8	36.3	969	6	AR477199	AR477199 Sequence
32	949.8	36.3	969	6	BD103571	BD103571 Novel G P
33	945	36.1	969	9	AF474988	AF474988 Homo sapi
34	938.6	35.9	969	6	AR477200	AR477200 Sequence
35	938.6	35.9	969	6	BD103572	BD103572 Novel G P
36	856.2	32.7	997	6	AX921831	AX921831 Sequence
37	837.8	32.0	969	6	E40003	E40003 Novel G pro
38	837.8	32.0	969	9	AF474989	AF474989 Homo sapi
39	836.2	31.9	969	6	AR477203	AR477203 Sequence
40	836.2	31.9	969	6	BD103575	BD103575 Novel G P
41	836.2	31.9	969	9	AF474991	AF474991 Homo sapi
42	834.6	31.9	969	6	E40000	E40000 Novel G pro
43	833	31.8	969	6	E39999	E39999 Novel G pro
44	833	31.8	969	6	AR477201	AR477201 Sequence
45	833	31.8	969	6	AX498194	AX498194 Sequence

ALIGNMENTS

RESULT 1	AX429465	Sequence 1 from Patent WO0234914.	2618 bp	DNA	linear	PAT 21-JUN-2002
LOCUS	AX429465					
DEFINITION	Sequence 1 from Patent WO0234914.					
ACCESSION	AX429465					
VERSION	AX429465.1	GI:21540763				
KEYWORDS						
SOURCE						
ORGANISM	Homo sapiens (human)					
REFERENCE	Zhao O., Beasley, E.M. and Wei, M.H.					
AUTHORS						
TITLE	Isolated human g-protein coupled receptors, nucleic acid molecules encoding human GPCR proteins, and uses thereof					
JOURNAL	Patent: WO 0234914-A 1 02-MAY-2002; PE CORP NY (US)					
FEATURES	Location/Qualifiers					
SOURCE	1..2618					
ORIGIN	/organism="Homo sapiens"					
	/mol_type="unassigned DNA"					
	/db_xref="taxon:9606"					
Query Match	100.0%;	Score 2618;	DB 6;	Length 2618;		
Best Local Similarity	100.0%;	Pred. No. 0;				
Matches 2618;	Conservative 0;	Mismatches 0;	Indels 0;	Gaps 0;		
Qy	1 AACAAATGCGCGAATTCGACAGATGAATCTAGTTGTTAAAGCGTGAACCT	60				
Db	1 AACAAATGCGCGAATTCGACAGATGAATCTAGTTGTTAAAGCGTGAACCT	60				
Qy	61 CTTCTCTCTCTCTTACTCTCTCTCTCTACCATGATGAGCGCTCGCTTTCCTTT	120				
Db	61 CTTCTCTCTCTCTTACTCTCTCTCTCTACCATGATGAGCGCTCGCTTTCCTTT	120				
Qy	121 CACGAGATTGGAAGCTTCCGAGAGCTCCCGACAGAGAGCGTGAAGCTTCTTGA	180				
Db	121 CACGAGATTGGAAGCTTCCGAGAGCTCCCGACAGAGAGCGTGAAGCTTCTTGA	180				
Qy	181 CAGTCTAGAGCATTAAGCAGTTAAACCATTTCTTCAATTAATTTCCAGTCTCAGG	240				
Db	181 CAGTCTAGAGCATTAAGCAGTTAAACCATTTCTTCAATTAATTTCCAGTCTCAGG	240				

Db 181 CAGTCTGATGAGCTATTAGCCAGTAAACCCATTCTTCATTAATTTCCAGTCTCAGG 240
 QY 241 TATTTCTTTTATGCAATTTGAGATGAATGAATCAATATACAGACAGAGAGGAGATGAGAA 300
 Db 241 TATTTCTTTTATGCAATTTGAGATGAATGAATCAATATACAGACAGAGAGGAGATGAGAA 300
 QY 301 ATCCCAAGGTGCTTTCTGCTGTCTTCCAGTCTCTGCTGTGTCTTCCAGTGTCTCAAT 360
 Db 301 ATCCCAAGGTGCTTTCTGCTGTCTTCCAGTCTCTGCTGTGTCTTCCAGTGTCTCAAT 360
 QY 361 TCCACCAAGAAACCAAAATTAATAAAGATCCCACTGATGTGTATCAATAGAACCACTCTCT 420
 Db 361 TCCACCAAGAAACCAAAATTAATAAAGATCCCACTGATGTGTATCAATAGAACCACTCTCT 420
 QY 421 TGGGATGTCAAAACAGATTAAGAAAGATGAAAGCAATCTCATGTGGTCAATCAGACTG 480
 Db 421 TGGGATGTCAAAACAGATTAAGAAAGATGAAAGCAATCTCATGTGGTCAATCAGACTG 480
 QY 481 GGTTCCTGAGCATGATTCACATCCAGTCTTGAGGTACAGAACTGACACCAATCAAG 540
 Db 481 GGTTCCTGAGCATGATTCACATCCAGTCTTGAGGTACAGAACTGACACCAATCAAG 540
 QY 541 GACGTGAGAGAGATCTCTTGTCTCAAGCAGAACCTTGAAGCTTCAAGGGGTGAGGTGATG 600
 Db 541 GACGTGAGAGAGATCTCTTGTCTCAAGCAGAACCTTGAAGCTTCAAGGGGTGAGGTGATG 600
 QY 601 TTTCCCTGTGCGCGCTGACAGAGAAACGCGTTGTGTCTGTGCTCTGCGGCTGCGCGATG 660
 Db 601 TTTCCCTGTGCGCGCTGACAGAGAAACGCGTTGTGTCTGTGCTCTGCGGCTGCGCGATG 660
 QY 661 GCAGAGACGCTGTCTTCATCTTACATCTTCAACCTGTGTGCGCGCGAATTCTTCTCTTA 720
 Db 661 GCAGAGACGCTGTCTTCATCTTACATCTTCAACCTGTGTGCGCGCGAATTCTTCTCTTA 720
 QY 721 GCGGCAATATATATGTTGCGCGGTAAACGGCTCATGAATATCGGCAATCCCATCTCCAAA 780
 Db 721 GCGGCAATATATATGTTGCGCGGTAAACGGCTCATGAATATCGGCAATCCCATCTCCAAA 780
 QY 781 TCTCATGCTGTGTATGACCTTTCCTTACTTATATAGGCTTAAGCATGTGAGGCGCATCA 840
 Db 781 TCTCATGCTGTGTATGACCTTTCCTTACTTATATAGGCTTAAGCATGTGAGGCGCATCA 840
 QY 841 GCACCGAGCGCTGCTGTCTTCATCTGTGTGCGCATCTGTATACCACTGCCCGCCAGAT 900
 Db 841 GCACCGAGCGCTGCTGTCTTCATCTGTGTGCGCATCTGTATACCACTGCCCGCCAGAT 900
 QY 901 ACCGTGATCGGTATGATGTGTCTGTGTGCGCGCATCTGTATACCACTGCCCGCCAGAT 960
 Db 901 ACCGTGATCGGTATGATGTGTCTGTGTGCGCGCATCTGTATACCACTGCCCGCCAGAT 960
 QY 961 AGTGAATGTCTGTATCTTCTGTGTATGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGT 1020
 Db 961 AGTGAATGTCTGTATCTTCTGTGTATGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGT 1020
 QY 1021 ATTTCAATTAACAATGCGGT 1080
 Db 1021 ATTTCAATTAACAATGCGGT 1080
 QY 1081 TCCGCTGT 1140
 Db 1081 TCCGCTGT 1140
 QY 1141 CCATCTCTCTCAAGATGT 1200
 Db 1141 CCATCTCTCTCAAGATGT 1200
 QY 1201 CCCTGT 1260
 Db 1201 CCCTGT 1260
 QY 1261 CCATTTTCTGT 1320
 Db 1261 CCATTTTCTGT 1320

QY 1321 CCTTAGGCAAGCGTCAAAATAGGCAAGAACTGAACTGTGTCTCCAGAGGCTCTGCAGG 1380
 Db 1321 CCTTAGGCAAGCGTCAAAATAGGCAAGAACTGAACTGTGTCTCCAGAGGCTCTGCAGG 1380
 QY 1381 ACAAGCCTGATGTGTATGAGGT 1440
 Db 1381 ACAAGCCTGATGTGTATGAGGT 1440
 QY 1441 GAAGCAGATTGAGAGAGTGAAGAGAACTGTGCGCTGTGAGACAGACTTGTGAGAGCA 1500
 Db 1441 GAAGCAGATTGAGAGAGTGAAGAGAACTGTGCGCTGTGAGACAGACTTGTGAGAGCA 1500
 QY 1501 TGCTGCGCTGCCACCTTGAACAATTATATGCAATTTTCTTACCTTGTGCTCAGAAATG 1560
 Db 1501 TGCTGCGCTGCCACCTTGAACAATTATATGCAATTTTCTTACCTTGTGCTCAGAAATG 1560
 QY 1561 TCTCATGTGTCTCTCAAGGTCTTTCGAATATGATGTGTATCTTAACTGAGAGTGTGAGTTT 1620
 Db 1561 TCTCATGTGTCTCTCAAGGTCTTTCGAATATGATGTGTATCTTAACTGAGAGTGTGAGTTT 1620
 QY 1621 CACCCATGGAAGCAATAGTGTGACAGTACAAATGTTTGGATTTCTCTGATATTAACAAT 1680
 Db 1621 CACCCATGGAAGCAATAGTGTGACAGTACAAATGTTTGGATTTCTCTGATATTAACAAT 1680
 QY 1681 ACAATTTCCCTGTATCTTGAACGTGAATCTTCTTACTGAAACATTTTCTGCACTTTTC 1740
 Db 1681 ACAATTTCCCTGTATCTTGAACGTGAATCTTCTTACTGAAACATTTTCTGCACTTTTC 1740
 QY 1741 ATTTGATTAATAAGAGGT 1800
 Db 1741 ATTTGATTAATAAGAGGT 1800
 QY 1801 CCTGATAGTATCAAAAGAGAGATTCCTTATTAATCTGTGACAGTATGTTCCCTGAGAAA 1860
 Db 1801 CCTGATAGTATCAAAAGAGAGATTCCTTATTAATCTGTGACAGTATGTTCCCTGAGAAA 1860
 QY 1861 TCATGTTCCCTTTATGACTGTGAGGCAATTAAGGCTGAGTGTGAAAGCTCAATCTTAAATAGT 1920
 Db 1861 TCATGTTCCCTTTATGACTGTGAGGCAATTAAGGCTGAGTGTGAAAGCTCAATCTTAAATAGT 1920
 QY 1921 GAGTTCTGTACTCTTAATTTCCATTTGAATTTCTGATATTAAGAAATTAATAGCCTTA 1980
 Db 1921 GAGTTCTGTACTCTTAATTTCCATTTGAATTTCTGATATTAAGAAATTAATAGCCTTA 1980
 QY 1981 GAGAGAGATTCCTCCTTCAATAAAGAGCTTGAAGATTTGTTTATGATAGACCTCTC 2040
 Db 1981 GAGAGAGATTCCTCCTTCAATAAAGAGCTTGAAGATTTGTTTATGATAGACCTCTC 2040
 QY 2041 CTGTCAATTTGTCAAGAGATGTGTGACATGTTGGCTTGTGTGTGTGTGTGTGTGTGTGT 2100
 Db 2041 CTGTCAATTTGTCAAGAGATGTGTGACATGTTGGCTTGTGTGTGTGTGTGTGTGTGTGT 2100
 QY 2101 GCGCCCTTCCCTTGAAGAACTGTATGTTCTTATTTAGCTTCCCTGAGCTAATAGAACTAG 2160
 Db 2101 GCGCCCTTCCCTTGAAGAACTGTATGTTCTTATTTAGCTTCCCTGAGCTAATAGAACTAG 2160
 QY 2161 TGAGAGAGCTTAAATATATGTTCCCAAGATTTTATTTTGTGCAATTTGGAACCTCAATATG 2220
 Db 2161 TGAGAGAGCTTAAATATATGTTCCCAAGATTTTATTTTGTGCAATTTTGTGCAATATG 2220
 QY 2221 ATTTTAAAGTGAATAATATCTTGAATAACCAATTAATTAATCAATCAATCTTCAATTT 2280
 Db 2221 ATTTTAAAGTGAATAATATCTTGAATAACCAATTAATTAATCAATCAATCTTCAATTT 2280
 QY 2281 GTAGAGAAATTTCTTCAATCTTCAAGTTTGTATTAATTTGTGTGTGTGTGTGTGTGTGT 2340
 Db 2281 GTAGAGAAATTTCTTCAATCTTCAAGTTTGTATTAATTTGTGTGTGTGTGTGTGTGTGT 2340
 QY 2341 TAGTTTATAGGCTTTTCAATGAGAGCAAACTGAAACATCTGACCTTTCATGACAA 2400
 Db 2341 TAGTTTATAGGCTTTTCAATGAGAGCAAACTGAAACATCTGACCTTTCATGACAA 2400

1482 GACAGACCTTGGAGAGCAATGCTGCCCTGACCCTTGAACATTAATATGACATTTTCTTA 1541
8487 GAGAGACCTTGGAGAGCAATGCTGCCCTGACCCTTGAACATTAATATGACATTTTCTTA 8428
1542 GCGCTTGCCTCAGAAATGCTCAGTGGTCCCTCAGAGCTTCTGAAATAGATGTTTATCTA 1601
8427 GCGCTTGCCTCAGAAATGCTCAGTGGTCCCTCAGAGCTTCTGAAATAGATGTTTATCTA 8368
1602 ACCTGACAGTTGGAGTTTTCACACGAGGAAGCATTAAGTCTGACAGACAAATGTTGGAT 1661
8367 ACCTGACAGTTGGAGTTTTCACACGAGGAAGCATTAAGTCTGACAGACAAATGTTGGAT 8308
1662 TCTCCTGATATTAACCAATACATTTTCCCTGTTATCTTGGACAGAACTTCTTCTACTGAA 1721
8307 TCTCCTGATATTAACCAATACATTTTCCCTGTTATCTTGGACAGAACTTCTTCTACTGAA 8248
1722 GACTTTTCTGACCTTTTCAATGTAATTAAGAGAGTTGCTGTCACAACTTAAACTCT 1781
8247 GACTTTTCTGACCTTTTCAATGTAATTAAGAGAGTTGCTGTCACAACTTAAACTCT 8188
1782 TCTTTATCTGTTTCTTCTACTGATATGATACAAAAGGAAGATTCCTTATTAATCTGTCA 1841
8187 TCTTTATCTGTTTCTTCTACTGATATGATACAAAAGGAAGATTCCTTATTAATCTGTCA 8128
1842 GACTATGTTCCCTGAGAAATCATGTTCCCTTTATGACTGAGGACATTAATGAGAGTTGA 1901
8127 GACTATGTTCCCTGAGAAATCATGTTCCCTTTATGACTGAGGACATTAATGAGAGTTGA 8068
1902 AGCTCAATCTTATTAATGAGAGTTGCTGCTAATCTTAAATTCATGTAATCTCAGATATA 1961
8067 AGCTCAATCTTATTAATGAGAGTTGCTGCTAATCTTAAATTCATGTAATCTCAGATATA 8008
1962 AGCAAAATTAATGAGAGTTGCTGCTAATCTTAAATTCATGTAATCTCAGATATA 2021
8007 AGCAAAATTAATGAGAGTTGCTGCTAATCTTAAATTCATGTAATCTCAGATATA 7948
2022 TTTTATGATAGAGCTTCTCCCTGATTTGTCACAGAGATGATGAGCTTGGCTTGGT 2081
7947 TTTTATGATAGAGCTTCTCCCTGATTTGTCACAGAGATGATGAGCTTGGCTTGGT 7888
2082 TCTAGTAAGACATGCTGCGCCCTTCCCTTGAAGACGTGTAATCTTATTAAGCTCT 2141
7887 TCTAGTAAGACATGCTGCGCCCTTCCCTTGAAGACGTGTAATCTTATTAAGCTCT 7828
2142 TCTGGAATATGATGATGAGAGAGCTTAAATATGCTCCACAGCTTCAATTTGGCC 2201
7827 TCTGGAATATGATGATGAGAGAGCTTAAATATGCTCCACAGCTTCAATTTGGCC 7768
2202 ATTGAAACCTCAATATGATTTTAAAGTGAATTAATCTTGAAGACATTTATTTATCA 2261
7767 ATTGAAACCTCAATATGATTTTAAAGTGAATTAATCTTGAAGACATTTATTTATCA 7708
2262 CTTCAGATATCTTTCAGTTGATGAGAGAAATCTTCAATCTTCAAGTTTGTATTAATTTGT 2321
7707 CTTCAGATATCTTTCAGTTGATGAGAGAAATCTTCAATCTTCAAGTTTGTATTAATTTGT 7648
2322 TCTGATTTGATCTTTCAGTTGATTTTAAAGCTGTTTACATGAGAGCAAACTGAAAAACA 2381
7647 TCTGATTTGATCTTTCAGTTGATTTTAAAGCTGTTTACATGAGAGCAAACTGAAAAACA 7588
2382 TCTGACCTTTCAGATGACATCTCAATATGATGATGATGATGATGATGATGATGATGATGAT 2441
7587 TCTGACCTTTCAGATGACATCTCAATATGATGATGATGATGATGATGATGATGATGATGAT 7528
2442 GAATTTGATATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 2501
7527 GAATTTGATATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 7468
2502 CCGAGAGCCCAAAATATTTGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 2561
7467 CCGAGAGCCCAAAATATTTGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 7408
2562 TGAACACAGAAATTAATTAATGAGAGATCTGCTGCTGAGAAAAA 2612

Db 7407 TGAACACAGAAATTAATTAATGAGAGATCTGCTGCTGAGAAAAA 7357
RESULT 3
AC090099/c
LOCUS
DEFINITION
ACCESSION
VERSION
KEYWORDS
SOURCE
ORGANISM
REFERENCE
AUTHORS
TITLE
JOURNAL
REFERENCE
AUTHORS
Birren, B., Linton, L., Nusbaum, C., Lander, E., Allen, N., Anderson, S.,
Barnes, N., Bastien, V., Boguslavsky, L., Bouckhelter, B., Brown, A.,
Camarata, J., Campopiano, A., Choepel, Y., Colangelo, M., Collins, S.,
Collymore, A., Cooke, P., Dearellano, K., Dewar, K., Diaz, J. S.,
Dodge, S., Faro, S., Ferreira, P., Fitzhugh, W., Gage, D., Galagan, J.,
Gardyna, S., Glende, S., Goyette, M., Graham, L., Grand-Pierre, N.,
Hagos, B., Heatford, A., Horton, L., Hulme, W., Iliev, I., Johnson, R.,
Jones, C., Karatas, A., Larocque, K., Lamazates, R., Landers, T.,
Lehoczy, J., Levine, R., Liu, G., Maclean, C., Macdonald, P.,
Marquis, N., Matthews, C., McCarthy, M., McEwan, P., McKernan, K.,
McPheeters, R., Meldrim, J., Meneus, L., Mihova, T., Mlenga, V.,
Murphy, T., Naylor, J., Nguyen, C., Norbu, C., Norman, C. H.,
O'Connor, T., O'Donnell, P., O'Neill, D., Oliver, J., Peterson, K.,
Punthang, P., Pierre, N., Pollara, V., Raymond, C., Retta, R.,
Rieback, M., Riley, R., Rise, C., Rogov, P., Roman, D., Roselli, M.,
Roy, A., Santos, R., Schauer, S., Schuback, R., Seaman, S., Severy, P.,
Sougnuez, C., Spencer, B., Stange-Thomann, N., Stojanovic, N.,
Strauss, N., Subramanian, A., Talamas, J., Testafaye, S., Theodore, J.,
Travers, M., Travis, N., Triggilio, J., Vassiliev, H., Viel, R., Vo, A.,
Wilson, B., Wu, X., Wyman, D., Ye, W. J., Young, G., Zainoun, J.,
Zembek, L., Zimmer, A. and Zody, M.
Direct Submission
Submitted (14-FEB-2001) Whitehead Institute/MIT Center for Genome
Research, 320 Charles Street, Cambridge, MA 02141, USA
3 (bases 1 to 172939)
Birren, B., Nusbaum, C., Lander, E., Ali, A., Allen, N., Anderson, S.,
Barnes, N., Bastien, V., Bloom, T., Boguslavsky, L., Bouckhelter, B.,
Camarata, J., Chang, J., Chazaro, B., Choepel, Y., Collymore, A.,
Cooke, A., Cooke, P., Dearellano, K., Dewar, K., Diaz, J. S., Dodge, S.,
Faro, S., Ferreira, P., FitzGerald, M., Gage, D., Galagan, J.,
Gardyna, S., Gord, S., Graham, L., Grand-Pierre, N., Hagos, B.,
Horton, L., Hulme, W., Iliev, I., Johnson, R., Jones, C., Kamat, A.,
Karatas, A., Kells, C., Landers, T., Levine, R., Lindblad-Toh, K.,
Liu, G., Maclean, C., Macdonald, P., Major, J., Matthews, C.,
McCarthy, M., Meldrim, J., Meneus, L., Mihova, T., Mlenga, V.,
Murphy, T., Naylor, J., Nguyen, C., Nicol, R., Norbu, C., Norman, C. H.,
O'Connor, T., O'Donnell, P., O'Neill, D., Oliver, J., Peterson, K.,
Punthang, P., Pierre, N., Raymond, C., Retta, R., Rise, C., Rogov, P.,
Roman, J., Roy, A., Schauer, S., Schuback, R., Seaman, S., Severy, P.,
Smith, C., Spencer, B., Stange-Thomann, N., Stojanovic, N., Talamas, J.,
Testafaye, S., Theodore, J., Topham, K., Travers, M., Vassiliev, H.,
Viel, R., Vo, A., Wilson, B., Wu, X., Wyman, D., Young, G., Zainoun, J.,
Zembek, L., Zimmer, A. and Zody, M.
Direct Submission
Submitted (05-SEP-2002) Whitehead Institute/MIT Center for Genome
Research, 320 Charles Street, Cambridge, MA 02141, USA
4 (bases 1 to 172939)
Birren, B., Nusbaum, C., Lander, E., Ali, A., Allen, N., Anderson, S.,
Barnes, N., Bastien, V., Bloom, T., Boguslavsky, L., Bouckhelter, B.,
Camarata, J., Chang, J., Chazaro, B., Choepel, Y., Collymore, A.,
Cooke, A., Cooke, P., Dearellano, K., Dewar, K., Diaz, J. S., Dodge, S.,
Faro, S., Ferreira, P., FitzGerald, M., Gage, D., Galagan, J.,
Gardyna, S., Gord, S., Graham, L., Grand-Pierre, N., Hagos, B.,

TITLE
JOURNAL
COMMENT

Horton, L., Hulme, W., Iliev, I., Johnson, R., Jones, C., Kamat, A., Karatas, A., Kells, C., Landers, T., Levine, R., Lindblad-Toh, K., Liu, G., Maclean, C., Macdonald, P., Major, J., Matthews, C., McCarthy, M., Meldrum, J., Menus, L., Mihova, T., Mlenga, V., Murphy, T., Naylor, J., Nguyen, C., Nicol, R., Norbu, C., Norman, C.H., O'Connor, T., O'Donnell, P., O'Neill, D., Oliver, J., Peterson, K., Phunkhang, P., Pierre, N., Raymond, C., Retta, R., Rise, C., Rogov, P., Roman, J., Roy, A., Schauer, S., Schuback, R., Seaman, S., Severy, P., Smith, C., Spencer, B., Strange-Thomann, N., Stojanovic, N., Talamas, J., Tesfaye, S., Theodore, J., Topham, K., Travers, M., Vassiliev, H., Viel, R., Vo, A., Wilson, B., Wu, X., Wyman, D., Young, G., Zainoun, J., Zembek, L., Zimmer, A. and Zody, M.

Direct Submission
Submitted (01-OCT-2002) Whitehead Institute/MIT Center for Genome Research, 320 Charles Street, Cambridge, MA 02141, USA
On Oct 1, 2002 this sequence version replaced gi:22726159.
All repeats were identified using RepeatMasker:
Smit, A.F.A. & Green, P. (1996-1997)
http://ftp.genome.washington.edu/RM/RepeatMasker.html

Genome Center
Center: Whitehead Institute/ MIT Center for Genome Research
Center code: WIRB
Web site: http://www-seq.wi.mit.edu
Contact: sequence_submissions@genome.wi.mit.edu
Project Information
Center project name: LI2764
Center clone name: 113_D_6

FEATURES
Source

Location/Qualifiers
1..172939
/organism="Homo sapiens"
/mol_type="genomic DNA"
/db_xref="taxon:9606"
/chromosome="11"
/map="11"
/clone="RP11-113D6"
/clone_11b="RP11-11 Human Male BAC"
complement(1..77)
/rpt_family="AluS"
complement(154..680)
/rpt_family="Char1el1"
complement(678..810)
/rpt_family="Char1el1"
complement(879..1057)
/rpt_family="L1M4"
complement(1395..1513)
/rpt_family="FLAM_A"
1870..1876
/note="<30 qual SNGL region"
1902..1910
/note="<30 qual SNGL region"
2672..2788
/rpt_family="5S"
complement(4009..4170)
/rpt_family="MIR"
complement(4210..4355)
/rpt_family="MERSB"
4356..4647
/rpt_family="AluSx"
complement(6554..6658)
/rpt_family="LTR41"
6916..6985
/rpt_family="MLT1C"
complement(7175..7359)
/rpt_family="LTR41"
complement(8308..8611)
/rpt_family="AluUb"
8622..9083
/rpt_family="MLT2B1"
complement(9265..10668)
/rpt_family="L1PB3"
10699..10809
/rpt_family="MER34"

repeat_region	10782..11264	/rpt_family="LTR48B"
repeat_region	complement(11344..11683)	/rpt_family="L2"
repeat_region	11835..12026	/rpt_family="L1MC4a"
repeat_region	12186..12660	/rpt_family="L1M2"
repeat_region	12879..13322	/rpt_family="L1MC4a"
repeat_region	13233..13522	/rpt_family="AluSx"
repeat_region	13523..13817	/rpt_family="L1MC4a"
repeat_region	14819..15108	/rpt_family="AluO"
repeat_region	15114..15175	/rpt_family="GAAA)n"
repeat_region	16150..16336	/rpt_family="MERSA"
repeat_region	complement(16587..16598)	/note="<30 qual SNGL region"
repeat_region	complement(16607..16612)	/note="<30 qual SNGL region"
repeat_region	complement(16746..17360)	/rpt_family="L1MB3"
repeat_region	17361..17664	/rpt_family="AluO"
repeat_region	complement(17665..17774)	/rpt_family="L1MB3"
repeat_region	complement(17827..17929)	/rpt_family="MIR"
repeat_region	18993..19298	/rpt_family="AluO"
repeat_region	19759..19821	/rpt_family="L2"
repeat_region	19933..19941	/note="<30 qual SNGL region"
repeat_region	20320..20609	/rpt_family="AluO"
repeat_region	complement(20684..20994)	/rpt_family="AluSx"
repeat_region	20713..20721	/note="<30 qual SNGL region"
repeat_region	20879..20883	/note="<30 qual SNGL region"
repeat_region	complement(21118..21398)	/rpt_family="AluSx"
repeat_region	21141..21147	/note="<30 qual SNGL region"
repeat_region	21144	/note="probably A, possibly G"
repeat_region	21146	/note="probably A, possibly G"
repeat_region	21153..21158	/note="<30 qual SNGL region"
repeat_region	21154	/note="probably T, possibly C"
repeat_region	21160..21167	/note="<30 qual SNGL region"

Query Match
Best Local Similarity 99.3%; Score 2128.6; DB 9; Length 172939;
Matches 2137; Conservative 0; Mismatches 14; Indels 0; Gaps 0;

QY	462	TCATGGTCATGACGCTGGGCTTTTGTGATGATTCATCAACATCCAGTCTTGAGTACA	521
DB	114002	TCACGGTCATGACGCTGGGCTTTTGTGATGATTCATCAACATCCAGTCTTGAGTACA	113943
QY	522	GAACGACACCAATCAAGGACGTGAGGAGACTCTTGCTCAAGACAGACCTTGAGCTTC	581
DB	113942	GAACGACACCAATCAAGGACGTGAGGAGACTCTTGCTCAAGACAGACCTTGAGCTTC	113883

ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Homidae; Homo.
REFERENCE 1 (bases 1 to 192618)
AUTHORS Waterston, R.H.
TITLE The sequence of Homo sapiens clone
JOURNAL Unpublished
AUTHORS 2 (bases 1 to 192618)
TITLE Waterston, R.H.
JOURNAL Direct Submission
Submitted (04-JUN-2000) Genome Sequencing Center, Washington
University School of Medicine, 4444 Forest Park Parkway, St. Louis,
MO 63108, USA
COMMENT On Jul 19, 2000 this sequence version replaced gi:9211362.
----- Genome Center -----
Center: Washington University Genome Sequencing Center
Center code: WUGSC
Web site: <http://genome.wustl.edu/gsc/index.shtml>
----- Project Information -----
Center project name: H NH0206C01
----- Summary Statistics -----
Sequencing vector: M13; 93%
Chemistry: Dye-primer ET; 93% of reads
Chemistry: Dye-terminator Big Dye; 7% of reads
Assembly program: Phrap; version 0.990319
Consensus quality: 179082 bases at least Q40
Consensus quality: 183120 bases at least Q30
Consensus quality: 185448 bases at least Q20
Insert size: 205000; agarose-fp
Insert size: 190318; sum-of-contrigs
Quality coverage: 3.99 in Q20 bases; agarose-fp
Quality coverage: 4.34 in Q20 bases; sum-of-contrigs

* NOTE: This is a 'working draft' sequence. It currently
* consists of 24 contrigs. The true order of the pieces
* is not known and their order in this sequence record is
* arbitrary. Gaps between the contrigs are represented as
* runs of N, but the exact sizes of the gaps are unknown.
* This record will be updated with the finished sequence
* as soon as it is available and the accession number will
* be preserved.
1 1279: contrig of 1279 bp in length
* 1280 1379: gap of unknown length
* 1380 3544: contrig of 2165 bp in length
* 3545 3644: gap of unknown length
* 3645 5441: contrig of 1797 bp in length
* 5442 8421: contrig of 2880 bp in length
* 8422 10096: contrig of 1575 bp in length
* 10097 10196: gap of unknown length
* 10197 11938: contrig of 1742 bp in length
* 11939 12038: gap of unknown length
* 12039 14708: contrig of 2670 bp in length
* 14709 14808: gap of unknown length
* 14809 17876: contrig of 3068 bp in length
* 17877 17976: gap of unknown length
* 17977 21534: contrig of 3558 bp in length
* 21535 21634: gap of unknown length
* 21635 24615: contrig of 2981 bp in length
* 24616 28301: contrig of 3586 bp in length
* 28302 28401: gap of unknown length
* 28402 35969: contrig of 7568 bp in length
* 35970 36069: gap of unknown length
* 36070 42473: contrig of 6404 bp in length
* 42474 42573: gap of unknown length
* 42574 49752: contrig of 7179 bp in length
* 49753 49852: gap of unknown length
* 49854 57604: contrig of 7752 bp in length
* 57605 57704: gap of unknown length

FEATURES
source
1. 192618
/organism="Homo sapiens"
/mol_type="genomic DNA"
/db_xref="taxon:9606"
/chromosome="20"
/clone="RP11-206C1"
1. 1279
/note="assembly_name:Contig19"
misc_feature
1380..3544
/note="assembly_name:Contig21"
misc_feature
3645..5441
/note="assembly_name:Contig22"
5442..8421
/note="assembly_name:Contig23"
8522..10096
/note="assembly_name:Contig24"
10197..11938
/note="assembly_name:Contig25"
12039..14708
/note="assembly_name:Contig26"
14809..17876
/note="assembly_name:Contig27"
17977..21534
/note="assembly_name:Contig28"
21635..24615
/note="assembly_name:Contig29"
24716..28301
/note="assembly_name:Contig30"
28402..35969
/note="assembly_name:Contig31
clone_end:SP6
vector_side:left"
36070..42473
/note="assembly_name:Contig32"
42574..49752
/note="assembly_name:Contig33"
49853..57604
/note="assembly_name:Contig34"
57705..67088
/note="assembly_name:Contig35"
67189..77656
/note="assembly_name:Contig36"
77757..92143
/note="assembly_name:Contig37"
92244..109240
/note="assembly_name:Contig38"
109341..124285
/note="assembly_name:Contig39"
124386..138608
/note="assembly_name:Contig40"
138709..153882
/note="assembly_name:Contig41"
153983..169994
/note="assembly_name:Contig42"
170095..192618
Location/Qualifiers
1. 192618
67088: contrig of 9384 bp in length
* 67089 67188: gap of unknown length
* 67189 77656: contrig of 10468 bp in length
* 77657 77756: gap of unknown length
* 77757 92143: contrig of 14387 bp in length
* 92144 92243: gap of unknown length
* 92244 109240: contrig of 16997 bp in length
* 109241 109341: gap of unknown length
* 109342 124285: contrig of 14945 bp in length
* 124286 124385: gap of unknown length
* 124386 138608: contrig of 14223 bp in length
* 138609 138708: gap of unknown length
* 138709 153882: contrig of 15174 bp in length
* 153883 153982: gap of unknown length
* 153983 169994: contrig of 16012 bp in length
* 169995 170094: gap of unknown length
* 170095 192618: contrig of 22524 bp in length.

ORIGIN		/note="assembly_name:Contig43"	
Query Match	75.9%; Score 1987.4; DB 2; Length 192618;	Best Local Similarity 96.4%; Pred. No. 0;	Matches 207; Conservative 0; Mismatches 71; Indels 7; Gaps 4;
Qy 462	TCATGGGTCATCAGACTGGGGTTCTGAGCATGAGATTCACCATCCAGTCTGGGTACA	521	
Db 156382	TCGAGGTCATCAGACTGGGGTTCTGAGCATGAGATTCACCATCCAGTCTGGGTACA	156441	
Qy 522	GAACCTGACCAATCAAAG	581	
Db 156442	GAACCTGACCAATCAAAG	156501	
Qy 582	ACGGGGCTGAGCTGATCGTTCCCTTGTCCGCTGACAGAGAAACGGGTGTCTCGG	641	
Db 156502	ACGGGGCTGAGCTGATCGTTCCCTTGTCCGCTGACAGAGAAACGGGTGTCTCGG	156561	
Qy 642	CTCCTGGGCTGCGGAG	701	
Db 156562	CTCCTGGGCTGCGGAG	156621	
Qy 702	GCCGACTTCTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCT	761	
Db 156622	GCCGACTTCTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCT	156681	
Qy 762	GCCGACTTCTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCT	821	
Db 156682	GCCGACTTCTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCT	156741	
Qy 822	AGCATGCTGAGCGGCTATGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG	881	
Db 156742	AGCATGCTGAGCGGCTATGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG	156801	
Qy 882	CACCTGCGCGCGCCAGATACCTGTCACTGTGTCTGTCTGTCTGTCTGTCTGTCTGTCC	941	
Db 156802	CACCTGCGCGCGCCAGATACCTGTCACTGTGTCTGTCTGTCTGTCTGTCTGTCTGTCC	156860	
Qy 942	CTGCTGCGGAGATCTGAGAGAGAGATCTGTCTGTCTGTCTGTCTGTCTGTCTGTCTGT	1001	
Db 156861	CTGCTGCGGAGATCTGAGAGAGAGATCTGTCTGTCTGTCTGTCTGTCTGTCTGTCTGT	156920	
Qy 1002	GTTTGTGTGAAAGTGAATTCATTCATTCATTCATTCATTCATTCATTCATTCATTCATTC	1061	
Db 156921	GTTTGTGTGAAAGTGAATTCATTCATTCATTCATTCATTCATTCATTCATTCATTCATTC	156980	
Qy 1062	CT--CTGTGGGTCTCAGCTGTCTGTCTGTCTGTCTGTCTGTCTGTCTGTCTGTCTGTCTGT	1118	
Db 156981	CTGTGGGTCTCAGCTGTCTGTCTGTCTGTCTGTCTGTCTGTCTGTCTGTCTGTCTGTCTGT	157040	
Qy 1119	CCGCTGACAGAGCT-GTACGTGACCATCTCTCTCAAGCTGTCTCTCTCTCTCTCTCTGTGG	1177	
Db 157041	CCGCTGACAGAGCTGTGACGTGACCATCTCTCTCAAGCTGTCTCTCTCTCTCTCTCTGTGG	157100	
Qy 1178	CTGCGCTTTTGGAGTGAAGTGGGCTTTTCCAGATTCACCTGATTTGGAAAGCTT	1237	
Db 157101	CTGCGCTTTTGGAGTGAAGTGGGCTTTTCCAGATTCACCTGATTTGGAAAGCTT	157158	
Qy 1238	ATTTTGTGATGTGATCATGTTTCTGCTGTCTGTCTGTCTGTCTGTCTGTCTGTCTGTCTGT	1297	
Db 157159	ATTTTGTGATGTGATCATGTTTCTGCTGTCTGTCTGTCTGTCTGTCTGTCTGTCTGTCTGT	157218	
Qy 1298	CATCATTTACTTCTTCTGTGGCTCTTTTGGAGAGCTGCAAAATAGGCAAACTGAAAGCT	1357	
Db 157219	CATCATTTACTTCTTCTGTGGCTCTTTTGGAGAGCTGCAAAATAGGCAAACTGAAAGCT	157278	
Qy 1358	GGTTTCTCAGAGGGCTCTGAGAGACAGCTGAGTGTGAATGAAGTGTGAAGGTGGCTTCC	1417	
Db 157279	GGTTTCTCAGAGGGCTCTGAGAGACAGCTGAGTGTGAATGAAGTGTGAAGGTGGCTTCC	157338	
Qy 1418	TCAGAAACCCTGAGCTGTCTGAGAGAGAGATTTGAGAGCTGAGAGAAACCTCTGCGCT	1477	
Db 157339	TCAGAAACCCTGAGCTGTCTGAGAGAGAGATTTGAGAGCTGAGAGAAACCTCTGCGCT	157398	
Qy 1478	GTGAGACAGACTTTTGAGAGCAATGCTGACCTGCACTTGTCAATTAATGATTTT	1537	
Db 157399	GTGAGACAGACTTTTGAGAGCAATGCTGACCTGCACTTGTCAATTAATGATTTT	157458	
Qy 1538	CTTACCTTCTGCTCAGAAATGTCTGAGTGTCTTCAAGGTCTTGTGAATGATTTT	1597	
Db 157459	CTTACCTTCTGCTCAGAAATGTCTGAGTGTCTTCAAGGTCTTGTGAATGATTTT	157518	
Qy 1598	TCCTACCTGACAGTGTGAGATTTTCACTGAGAGAGAGATTTGCTGACAGTCAATGTTT	1657	
Db 157519	TCCTACCTGACAGTGTGAGATTTTCACTGAGAGAGAGATTTGCTGACAGTCAATGTTT	157578	
Qy 1658	GGATTCCTCTGATTAATCAATCAATTTTCCGTTATCTGTGACGATCTTCTCTAC	1717	
Db 157579	GGATTCCTCTGATTAATCAATCAATTTTCCGTTATCTGTGACGATCTTCTCTAC	157638	
Qy 1718	TGAACCTTTTCTGACCTTTTCTGATTTTAAAGAGAGTGTGCTCACACCTTAAA	1777	
Db 157639	TGAACCTTTTCTGACCTTTTCTGATTTTAAAGAGAGTGTGCTCACACCTTAAA	157698	
Qy 1778	CTTCTTTTACTTGTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCT	1837	
Db 157699	CTTCTTTTACTTGTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCT	157758	
Qy 1838	GTGAGACTAATGTTCCCTGAAATCAATGTTCCCTTTTATGACGTGAGAGCTTCTGAGT	1897	
Db 157759	GTGAGACTAATGTTCCCTGAAATCAATGTTCCCTTTTATGACGTGAGAGCTTCTGAGT	157818	
Qy 1898	TGGAAGCTCAATCTTATTAAGTGTGCTGTCTTCACTTCAATTTCCATTAATCTCAGA	1957	
Db 157819	TGGAAGCTCAATCTTATTAAGTGTGCTGTCTTCACTTCAATTTCCATTAATCTCAGA	157878	
Qy 1958	TATTAACCAAAATATGACCTTGAAGAGAGATTTCTCTTCAATAAACAGCTTGAAGA	2017	
Db 157879	TATTAACCAAAATATGACCTTGAAGAGAGATTTCTCTTCAATAAACAGCTTGAAGA	157938	
Qy 2018	TTGGTTTATGAATGAGGCTCTCTGTCTTGTGCAACACATGAGTGAAGTGGGCTT	2077	
Db 157939	TTGGTTTATGAATGAGGCTCTCTGTCTTGTGCAACACATGAGTGAAGTGGGCTT	157998	
Qy 2078	GGTTTCTAGTAAAGACATGAGGCTTCCCTGAGAGAGCTGTGAATCTTATTTAG	2137	
Db 157999	GGTTTCTAGTAAAGACATGAGGCTTCCCTGAGAGAGCTGTGAATCTTATTTAG	158058	
Qy 2138	CTCTTCTGAGACTAATGACCTGAGAGAGCTTAAATATGTCACCAAGTTTCAATTT	2197	
Db 158059	CTCTTCTGAGACTAATGACCTGAGAGAGCTTAAATATGTCACCAAGTTTCAATTT	158118	
Qy 2198	GCCATTTGAAAGCTCAATTTGATTTTAAAGTGAATTTATCTTGAACCAATTTATTA	2257	
Db 158119	GCCATTTGAAAGCTCAATTTGATTTTAAAGTGAATTTATCTTGAACCAATTTATTA	158178	
Qy 2258	TTGACTTGAAGATCTTCTGAGTGTGAGAGAAATCTTCACTTCCAGTTTGTATPAA	2317	
Db 158179	TTGACTTGAAGATCTTCTGAGTGTGAGAGAAATCTTCACTTCCAGTTTGTATPAA	158238	
Qy 2318	TTGTTCTGATTTGAATCTTCAATTAATGATTTTATGAGTGTGATTAATGAGAGCAAACTGAA	2377	
Db 158239	TTGTTCTGATTTGAATCTTCAATTAATGATTTTATGAGTGTGATTAATGAGAGCAAACTGAA	158298	
Qy 2378	AACATCTGACCTTCTTCAATGACATCTCAATTAATGATTTGAGTAAATCTTACAGTTGG	2437	
Db 158299	AACATCTGACCTTCTTCAATGACATCTCAATTAATGATTTGAGTAAATCTTACAGTTGG	158358	
Qy 2438	TACAGATTTCTGATTAATGCTGTGACATCAATGAACCTGAAATTTGTCTAAGGAAA	2497	
Db 158359	TACAGATTTCTGATTAATGCTGTGACATCAATGAACCTGAAATTTGTCTAAGGAAA	158418	
Qy 2498	TAAAGCAGAGCGCAACATATTTGATTTCAATTTCTATGAGTATCAAAATTTAGGAAA	2557	
Db 158419	TAAAGCAGAGCGCAACATATTTGATTTCAATTTCTATGAGTATCAAAATTTAGGAAA	158478	

[illegible]

OY	1662	TCTCCGTGATATTAACAATAACATTTTCCCGGTATCTGGAACGAAATCTTCCACTGAA	1721
Db	154885	TCTCCTTGATATTACCAATACATTTTCCCTGTTATCTTGACACTGAATCTTCTACTGAA	1548286
OY	1722	CACCTTTTCTGCACTTTTCATGTGTAATAAAGAGTGTGTCGACACCCCTAAACTCT	1781
Db	154825	CACCTTTTCTGCACTTTTCATGTGTAATAAAGAGTGTGTCGACACCCCTAAACTCT	1547666
OY	1782	TCTTTATACTTGTGTTCTACCTGATAGTATCAAAAAGGAAGATTCTTATTATCTGTCA	1841
Db	154765	TCTTTATACTTGTGTTCTACCTGATAGTATCAAAAAGGAAGATTCTTATTATCTGTCA	1547066
OY	1842	GACATGTTCCCTCGAAAAATCAANGTCCCTTTATAGCTGGAGGCAATTAATGAGTTGGA	1901
Db	154705	GACATGTTCCCTCGAAAAATCAANGTCCCTTTATAGCTGGAGGCAATTAATGAGTTGGA	1546466
OY	1902	AGCTCAATTCTTAATAAGTAGAGTTCTGTCACTCTAAATTCATGTAATCTCAGATATA	1961
Db	154645	AGCTCAATTCTTAATAAGTAGAGTTCTGTCACTCTAAATTCATGTAATCTCAGATATA	1545866
OY	1962	AAGCAAAATATGACCTTGTGAGAGAGATTCTCCCTCATATAAACAGTCTTAGAAATTGG	2021
Db	154585	AAGCAAAATATGACCTTGTGAGAGAGATTCTCCCTCATATAAACAGTCTTAGAAATTGG	1545266
OY	2022	TTTATGAAATAGCCCTCTCCTGTCAATTGTGCAAGATGTTGACATGTTGGCCTGGTT	2081
Db	154525	TTTATGAAATAGCCCTCTCCTGTCAATTGTGCACA-CATGTGACACAGTTGGCCTGGTT	1544676
OY	2082	TCTAGTAAAGACATGCTGGGCCCTTCCCTTGAGAACTGTGTAAGTCTTATTTAGCTCT	2141
Db	154466	TCTAGTAAAGACATGCTGGGCCCTTCCCTTGAGAACTGTGTAAGTCTTATTTAGCTCT	1544076
OY	2142	TCTGTGACTAATGAACTAGTAGAGAGCCTTAATAATGTCGCCACAGTTTCATTTGGCC	2201
Db	154406	TCTGTGACTAATGAGACCAAGTAGAGAGCCTTAATAATCACCCCAAGTTCCATTTGGGCC	1543476
OY	2202	ATTGGAACCTCAATATGATTTTAAAGGGAATTAATCTGTAAGAACCATTTATTTATCA	2261
Db	154346	ATTGGAACCTCAATATGATTTTAAAGGGAATTAATCTGTAAGAACCATTTATTTATCA	1542896
OY	2262	CTTACAGATCTCTTTCAGTTGTAGAGAAATCTTCATACTTCCAGTGTGTTGTATAAATTGT	2321
Db	154288	TTTACAGAGTCTTTAAGTTGTAGAGAAATCTTCATACTTTCAGTGTGTTGTATAAATTGT	1542229
OY	2322	TCTGATTTGTAACCTTTCAGTTAGTTTATGCTGTTTACATGAGAGCAAAACTGAAAACA	2381
Db	154228	TCTGCTGTGTAACCTTTCAGTTAATTTTATGCTGTTTAACTAAGAGCAAAACTGAAAACA	1541696
OY	2382	TCTGACCTTTCAGATGCAATCTCAATTATGTAATCTGTAATTAACCTTTCAGTTGGTACA	2441
Db	154168	TCTGACTTTTCAGATGCTAATGTAATTAATTAATGTAATTAACCTTTCAGTTGGTACA	1541096
OY	2442	GAAATCTGATATGCTGTGACATATACATGAACTGTGAAATATTGTGCTAAGAAAAATAG	2501
Db	154108	AAATCTGATATGCTGTGACTGTGACTTAATATGAAACAGGAAACATGTGTGAAGAAAAATAG	1540496
OY	2502	CCAGACGCCA---AACATATTGTAAAGTTCAAAATCTATGAGGTATCCAAATTAAGAAAT	2558
Db	154048	CCAGATACCCAAAGAACAAATATTGTATATGATCAAAATTTCAACAGGTATCCAAATTAAGAAAT	1539896
OY	2559	TCTTGAACACAGAAAAATTAATTTAGAGGAAATCCGTGGTCTGGAAGAAAAAATAAATAA	2612
Db	153988	TCTGAACACAGAAAAATTAATTTAGAGGAAATCCGTGGTCTGGAATGAAAGAA 153935	

RESULT 6	
AC023078/c	
LOCUS	AC023078
DEFINITION	Homo sapiens chromosome 11, clone RP11-583F24, complete sequence.
ACCESSION	AC023078
VERSION	AC023078.9
KEYWORDS	GI:20198710
HTG.	

SOURCE	Homo sapiens (human)
ORGANISM	Homo sapiens
	Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE	1 (bases 1 to 163718)
AUTHORS	Bairren,B., Linton,L., Nusbaum,C. and Lander,E.
TITLE	Homo sapiens chromosome 11, clone RP11-583P24
JOURNAL	Unpublished
REFERENCE	2 (bases 1 to 163718)
AUTHORS	Bairren,B., Linton,L., Nusbaum,C., Lander,E., Abraham,H., Allen,N., Binkley,J., Blakesley,R., Bork,P., Brown,D., Cohen,M., Cook, D., Coulson,A., Duggan,G., Ewan,T., Gibson,T., Gillis,J., Har rison,W., Holt,I., Hughes,S., Hunter,D., Jones,K., Kuan, P., MacKenzie,A., Marshall,A., McQuibban,G., Mortlock,A., P earce,N., Perkins,D., Platts,R., Rago,G., Searles,C., Shownk en,K., Skuse,D.H., Smith,M., Steward,C., Sutherland,I., Tan gredi,V.P., Young,I.

Landers, T., Ienhoczky, J., Levine, R., Irieu, C., Liu, G., Locke, K., MacDonald, P., Margulis, N., McEwan, P., McGuck, A., McInernan, K., McPheters, R., Meldrum, J., Meunus, L., Morrow, J., Naylor, J., Norman, C.H., O'Connor, T., O'Donnell, P., Oliver, T.M., Peterson, K., Pierre, N., Pisan, C., Pollara, V., Raymond, C., Riley, R., Rothman, D., Roy, A., Santos, R., Severy, P., Spencer, B., Strange-Thomann, N., Stojanovic, N., Subramanian, A., Talmag, J., Tesfaye, S., Theodore, J., Titzell, A., Vassiliev, H., Viel, R., Wu, A., Wu, X., Wyman, D., Ye, W., J. Zimmerman, D. and Zody, W.

JOURNAL Submitted (08-FEB-2000) Whitehead Institute/MIT Center for Genome

REFERENCE

AUTHORS

AUTHORS
 Birren, B., Linton, L., Nussbaum, C., Lander, E., Ali, A., Allen, N.,
 Anderson, S., Barna, N., Bastien, V., Bloom, T., Boguslavsky, L.,
 Boukhalter, B., Brown, A., Camarata, J., Campopiano, A., Chang, J.,
 Chazaro, B., Choegel, Y., Colangelo, M., Collins, S., Collymore, A.,
 Cook, A., Cooke, P., Deatellano, K., Dewar, K., Diaz, J. S., Dodge, S.,
 Fairo, S., Ferris, P., Fitzhugh, W., Gage, D., Galagan, J., Gardyna, S.,
 Gird, S., Gird, S., Goyette, M., Graham, L., Grand-Pierre, N.,
 Hagos, B., Horton, L., Hulme, W., Iliev, I., Johnson, R., Jones, C.,
 Kama, A., Karatas, A., Kells, C., Lacroque, K., Lamazares, R.,
 Landers, T., Lenczky, J., Levine, R., Lindblad-Toh, K., Liu, G.,
 Maclean, C., MacDonald, P., Major, J., Margulis, N., Matthews, C.,
 McCaithy, M., McEwan, P., McKernan, K., Meldrum, J., Menzies, L.,
 Mihova, T., Mlenga, V., Murphy, T., Naylor, J., Nguyen, C., Nicol, R.,
 Norbu, C., Norman, C. H., O'Connor, T., O'Donnell, P., O'Neill, D.,
 Oliver, J., Peterson, K., Phunthang, P., Pierre, N., Pollara, V.,
 Raymond, C., Retta, R., Rieback, W., Riley, R., Riese, C., Rogov, P.,
 Roman, J., Rosetti, M., Roy, A., Santos, R., Schauer, S., Schnippsack, R.,
 Seaman, S., Severy, P., Spencer, B., Stange-Thomann, N., Stojanovic, N.,
 Strauss, N., Subramanian, A., Talamas, J., Tesfaye, S., Theodore, J.,
 Topham, K., Travers, M., Travis, N., Trigglio, J., Vassiliev, H.,
 Viel, R., Vo, A., Wilson, B., Wu, X., Wyman, D., Ye, W. J., Young, G.,
 Zaioun, R., Zembek, L., Zimmer, A. and Zody, M.

TITLE

JOURNAL

REFERENCE

AUTHORS

.....

QY 762 CGCCATCCCATCTCCAAAATCTCTGATGCTGATGACCTTTCCTACTTTATAGCCCTTA 821
 DB 71267 CCCCATACATCTCTAAAATCTCTATCTGATGATGTTTTCCTACTTTGAGGCCCTG 71208
 QY 822 AGCATGCTGAGGCCATCAGACCCGAGCCCTGCTGCTCACTCTGAGGCCATCTGATAC 881
 DB 71207 AGCTTTCGATGAGCCGTGAGCACCGAGCCCTGCTGCTGCTGAGGCCATCTGATAC 71148
 QY 882 CACTGCGCGCGCCGAGATACCTGATGCTGATGCTGATGCTGCTGAGGCCCTGCTGCTC 941
 DB 71147 CGCTGCGCGCGCCGAGATACCTGATGCTGATGCTGATGCTGCTGAGGCCCTGCTGCTC 71088
 QY 942 CTGCTGCGCGAGATCTCTGAGATGATGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 1001
 DB 71087 CTGCTGCGCGAGATCTCTGAGATGATGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 71028
 QY 1002 GTTGTGCTGAAACCTCAGATTTCAATACATGCGCTGCTGCTGCTGCTGCTGCTGCTGCT 1061
 DB 71027 GCTTGTGCTGAAACCTCAGATTTCAATACATGCGCTGCTGCTGCTGCTGCTGCTGCTGCT 70968
 QY 1062 CTCTGCGCGCTCAGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 1121
 DB 70967 CTCTGCGCGCTCAGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 70908
 QY 1122 CTGACCAAGCTGATGCTGATCCTCTCTCAGAGTCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 1181
 DB 70907 CTGACCAAGCTGATGCTGATCCTCTCTCAGAGTCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 70848
 QY 1182 CCTTTGGCATTCAGTGGCCCTGTTTTCAGAGATCCACTGATGGAAGTCTTATTT 1241
 DB 70847 CCTTTGGCATTCAGTGGCCCTGTTTTCAGAGATCCACTGATGGAAGTCTTATTT 70788
 QY 1242 TGTGATGTCATCTGATTTTCCATTTTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 1301
 DB 70787 TGTGATGTCATCTGATTTTCCATTTTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 70728
 QY 1302 ATTACTTCTTCTGCTGCTGCTTCTTGAAGCAGCTCAAAATAGCAGAACTGTAAGCTGTT 1361
 DB 70727 ATTACTTCTTCTGCTGCTGCTTCTTGAAGCAGCTCAAAATAGCAGAACTGTAAGCTGTT 70668
 QY 1362 CTCGAGAGGCTCTGACAGACGCTGCTGATGATGAAGTGTGAGAGGCTGCTGCTGCTGCTGCT 1421
 DB 70667 CTCGAGAGGCTCTGACAGACGCTGCTGATGATGAAGTGTGAGAGGCTGCTGCTGCTGCTGCT 70608
 QY 1422 GAACCCCTGAGCTGCTGAGGAGAGATGAGAGAGTGAAGAACTCTGCTGCTGCTGCTGCTGCTGCT 1481
 DB 70607 GAACCCCTGAGCTGCTGAGGAGAGATGAGAGAGTGAAGAACTCTGCTGCTGCTGCTGCTGCTGCT 70548
 QY 1482 GACAGACTTGAAGAGCAATGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 1541
 DB 70547 GACAGACTTGAAGAGCAATGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 70488
 QY 1542 GCTTCTGCTGCTGAGAAATGCTCAGTGTCTCCTCAAGTCTTCTGAAATGATGTTATCTTA 1601
 DB 70487 GCTTCTGCTGCTGAGAAATGCTCAGTGTCTCCTCAAGTCTTCTGAAATGATGTTATCTTA 70428
 QY 1602 ACCGAGAGTGAAGTCTTCAACCATGGAAGCAATTAAGTCTGACGTAAGATGTTGAT 1661
 DB 70427 ACCGAGAGTGAAGTCTTCAACCATGGAAGCAATTAAGTCTGACGTAAGATGTTGAT 70368
 QY 1662 TCTCTGATATTAACAATACATTTTCTGATCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 1721
 DB 70367 TCTCTGATATTAACAATACATTTTCTGATCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 70308
 QY 1722 CACTTTTCTGACATTTTCTGATGATTAATAAGAGTGTCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 1781
 DB 70307 CACTTTTCTGACATTTTCTGATGATTAATAAGAGTGTCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 70249
 QY 1782 TCTTTATCTGTTCTTCTACCGATGATGATTAATAAGAGTGTCTTCTTATTAATCTGCTA 1841
 DB 70248 TCTTTATCTGTTCTTCTACCGATGATGATTAATAAGAGTGTCTTATTAATCTGCTA 70189

QY 1842 GACTATGTTCCCTGAGAAATCATGTTCCCTTTTATGACTGAGGCAATTAATCTGACGTTGGA 1901
 DB 70188 GACTATGTTCCCTGAGAAATCATGTTCCCTTTTATGACTGAGGCAATTAATCTGACGTTGGA 70129
 QY 1902 AGCTCAATTTCTTAATGAAGTCTGCTACTCTGCTCAATTAATGATTAATCTGATGATTA 1961
 DB 70128 AGCTCAATTTCTTAATGAAGTCTGCTACTCTGCTCAATTAATGATTAATCTGATGATTA 70069
 QY 1962 AAGCAAAATTAATGACCTTAGAGAGATCTCTCTCAATAAAAACAGCTTAGAAATTTG 2021
 DB 70069 AAGCAAAATTAATGACCTTAGAGAGATCTCTCTCAATAAAAACAGCTTAGAAATTTG 70009
 QY 2022 TTTTATGAATACCTCTCTCTGCTCAATTTGCTCAAGATGCTGATGCTGCTGCTGCTGCT 2081
 DB 70008 TTTTATGAATACCTCTCTCTGCTCAATTTGCTCAAGATGCTGATGCTGCTGCTGCTGCT 69949
 QY 2082 TCTAGTAAAGCAATGCGGCCCTTCCCTGAGAACTGCTGATGCTTATTAAGCTT 2141
 DB 69948 TCTAGTAAAGCAATGCGGCCCTTCCCTGAGAACTGCTGATGCTTATTAAGCTT 69889
 QY 2142 TCTGAGCTAAATGAATGAGAGGAGGCTTAAATATGCTCCACAGTTTCAATTTGAGCC 2201
 DB 69888 TCTGAGCTAAATGAATGAGAGGAGGCTTAAATATGCTCCACAGTTTCAATTTGAGCC 69832
 QY 2202 ATTGAAACCTCAATATTTGATTTTAAAGTGAATTAATCTTGAATACATTTATTTCA 2261
 DB 69831 ATTGAAACCTCAATATTTGATTTTAAAGTGAATTAATCTTGAATACATTTATTTCA 69772
 QY 2262 CTGACAGTCTTCAATGATGAGAGATTTCTCAATCTTCAAGTTTGTATTAATTTGT 2331
 DB 69771 CTGACAGTCTTCAATGATGAGAGATTTCTCAATCTTCAAGTTTGTATTAATTTGT 69712
 QY 2332 TCTGATGTTAACTTCAATGATGATTTTATGCTGTTTATGATGAGAGCAAACTGAAAA 2381
 DB 69711 TCTGATGTTAACTTCAATGATGATTTTATGCTGTTTATGATGAGAGCAAACTGAAAA 69652
 QY 2382 TCTGATGTTAACTTCAATGATGATTTTATGCTGTTTATGATGAGAGCAAACTGAAAA 2441
 DB 69651 TCTGATGTTAACTTCAATGATGATTTTATGCTGTTTATGATGAGAGCAAACTGAAAA 69592
 QY 2442 GAATTCGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 2501
 DB 69591 GAATTCGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 69533
 QY 2502 CCAGAGGCGCAAAATATTTGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 2561
 DB 69532 CCAGAGGCGCAAAATATTTGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 69475
 QY 2562 TGAACACAGAAATTAATTTAGAGAGATCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 2600
 DB 69474 TGAACACAGAAATTAATTTAGAGAGATCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 69436

RESULT 7
 AC139482/c 165490 bp DNA linear HTG 04-FEB-2003
 LOCUS Home sapiens chromosome 5 clone RP11-583F24, WORKING DRAFT
 DEFINITION
 ACCESSION AC139482.1 GI:28201512
 VERSION HTG; HTGS_PHASE1; HTGS_DRAFT; HTGS_ACTIVEPIN.
 KEYWORDS Homo sapiens (human)
 SOURCE
 ORGANISM Homo sapiens
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.
 REFERENCE
 AUTHORS DOE Joint Genome Institute.
 JOURNAL Sequencing of Human Chromosome 5
 TITLE Unpublished
 DOB Joint Genome Institute.
 REFERENCE
 AUTHORS Direct Submission
 JOURNAL Submitted (04-FEB-2003) Production Sequencing Facility, DOE Joint

COMMENT

Genome Institute, 2800 Mitchell Drive, Walnut Creek, CA 94598, USA
-----Genome Center
Center: Joint Genome Institute
Center Code: JGI
Web site: <http://www.jgi.doe.gov>

Project Information
Center Project Name: 1518520
Center clone name: RPCI-11_583F24

Summary Statistics

Consensus quality: 147035 bases at least Q40
Consensus quality: 155210 bases at least Q30
Consensus quality: 158567 bases at least Q20
Estimated insert size: 175000; agarose-fp estimation
Estimated insert size: 164390; sum-of-contigs estimation
Quality coverage: 3.94 in Q20 bases; agarose-fp estimation
Quality coverage: 4.19 in Q20 bases; sum-of-contigs estimation.
* NOTE: This is a 'working draft' sequence. It currently
* consists of 12 contigs. The true order of the pieces
* is not known and their order in this sequence record is
* arbitrary. Gaps between the contigs are represented as
* runs of N, but the exact sizes of the gaps are unknown.
* This record will be updated with the finished sequence
* as soon as it is available and the accession number will
* be preserved.

1 1249: contig of 1249 bp in length
* 1250 1349: gap of unknown length
* 1350 3680: contig of 2331 bp in length
* 3681 3780: gap of unknown length
* 3781 6668: contig of 2888 bp in length
* 6669 6768: gap of unknown length
* 6769 10866: contig of 4098 bp in length
* 10867 10966: gap of unknown length
* 10967 17019: contig of 6053 bp in length
* 17020 17119: gap of unknown length
* 17120 22433: contig of 5314 bp in length
* 22434 22533: gap of unknown length
* 22534 34645: contig of 12112 bp in length
* 34646 34745: gap of unknown length
* 34746 49173: contig of 14428 bp in length
* 49174 49273: gap of unknown length
* 49274 67058: contig of 17785 bp in length
* 67059 67158: gap of unknown length
* 67159 84928: contig of 17770 bp in length
* 84929 85028: gap of unknown length
* 85029 114464: contig of 29436 bp in length
* 114465 114564: gap of unknown length
* 114565 165490: contig of 50926 bp in length.

FEATURES

source

1.165490
/organism="Homo sapiens"
/mol_type="genomic DNA"
/db_xref="taxon:9606"
/chromosome="5"
/clone="RP11-583F24"
/clone_11p="RPCI human BAC library 11"

ORIGIN

Query Match 68.5%; Score 1792.2; DB 2; Length 165490;
Best Local Similarity 91.1%; Pred. No. 0;
Matches 1949; Conservative 0; Mismatches 183; Indels 7; Gaps 4;

QY 462 TCATGGTCATCAGATCGGGTTTCTGAGCATGATTCACCATCCAGCTTGGGTACA 521
DB 74839 TCCAGGGTTCACACAGATGAGGTTTCTGAGCATGATTCACCATCCAGCTTGGGTACA 74780
QY 522 GAATGACACCAATCAACGAGCATGAGAGCATCTCTGCTACAGCAGACCTGAGCTTC 581
DB 74779 GAATGACACCAATCAACGAGCATGAGAGCATCTCTGCTACAGCAGACCTGAGCTTC 74720
QY 582 ACGGGGTCGACGTGATCTGTTCCCTTGTCCGCGTCGACAGAAACGCGGTGTGCTCTGG 641
|||||

DB 74719 ACGGTGTCGACGTGATCGTTTCCCTTGTCCGCGTCGACAGAAACGAGTTGCTCTGG 74660
QY 642 CTCCTGGGTCGCGCATGCGCAGAGAAAGCTGTCTCATCTACATCTCAACCTGTCGCG 701
DB 74659 CTCCTGGGTCGCGCATGCGCAGAGAAAGCTGTCTCATCTACATCTCAACCTGTCGCG 74600
QY 702 GCCGACTTCTCTCTCTTAAAGCGGCAATTATATGTTGCGCGTTACGCCCTCATCAATATC 761
DB 74599 GCACATCTCTCTCTCTCTCAAGGCGGCTTATATATTCCTGTTAACTTCACTCATATC 74540
QY 762 CGCCATCCCATCTCAAAATCTCTAGTCTGTGATGACCTTTCCCTTATATAGGCTTA 821
DB 74539 CCCCATCACTCTCAAAATCTCTATCTGTGATGATGTTTCTTCTTTCAGGCGCTG 74480
QY 822 AGCATGTCGAGCGCATCAGACCGGCGCTGTCAATCTGAGGCCCATCTGAGTAC 881
DB 74479 AGCTTCTGAGTGCCTGAGACCGAGCGGCTGTGCGCTCTGTGAGCCATCTGATAC 74420
QY 882 CACTGCGCGCGCCCGCAGATACCTGTCAATGAGTGTGCTGTGCGGCGCTGTCC 941
DB 74419 CGCTGCCACCGCCCGCAGACCTGTCAAGCGGTGTGTCTGTGCGGCGCTGTCC 74360
QY 942 CTGCTGCGAGATATCTGAGTGTGATTTCTGTGATCTTCTGTTATGTCGTGATTTCT 1001
DB 74359 CTGCTGCGAGCATCTGAGTGTGATTTGTCGCTTCTGTCAGTGTGCTGATTTCT 74300
QY 1002 GTTGTGTCGAAACGTGATTTCTATTAATGCGCGGCTTTTATATGATGATTT 1061
DB 74299 GCTTGTGTCGAAACGTGATTTCTATTAATGCGCGGCTTTTATATGATGATTT 74240
QY 1062 CTCCTGGGTCAGCTGCTGTCTGTGATGAGATTTCTGTGATGATCCCGAAGATCCG 1121
DB 74239 CTCCTGGGTCAGCTGCTGTCTGTGATGAGATTTCTGTGATGATCCCGAAGATCCG 74180
QY 1122 CTGACAGGCTGTGATGATGATCTCTCTCAAGTGTCTCTCTCTCTGTGCGCTG 1181
DB 74179 CTGACAGGCTGTGATGATGATCTCTCTCAAGTGTCTCTCTCTCTGTGCGCTG 74120
QY 1182 CCCCTTGGCATTCAGTGTGCGCTGTCTCTCAAGTGTCTCTCTCTCTGTGCGCTG 1241
DB 74119 CCCCTTGGCATTCAGTGTGCGCTGTCTCTCAAGTGTCTCTCTCTCTGTGCGCTG 74060
QY 1242 TGTGATGATCTGATGATTTCCATTTTCTGTCTCTCTCAAGTGTCTCTCTCTGTGCGCTG 1301
DB 74059 TGTGATGATCTGATGATTTTCTATTTTCTGTCTCTCTCAAGTGTCTCTCTCTGTGCGCTG 74000
QY 1302 ATTACTTCTGTGCGCTCTCTTAAAGCAGCTCAAAATAGGCAAACTTGAAGCTGTT 1361
DB 73999 ATTACTTCTGTGCGCTCTCTTAAAGCAGCTCAAAATAGGCAAACTTGAAGCTGTT 73940
QY 1362 CTCAGAGGCTCTGAGAGCAGCGCTGAGATGATGATGATGATGATGATGATGATGAT 1421
DB 73939 CTCAGAGGCTCTGAGAGCAGCGCTGAGATGATGATGATGATGATGATGATGATGAT 73880
QY 1422 GAAACCTGAGAGCTGTGAGAGAGATGAGAGATGAGAGAGATGAGAGAGATGAGAGAT 1481
DB 73879 GAAATCTGAGAGCTGTGAGAGAGATGAGAGATGAGAGAGATGAGAGAGATGAGAGAT 73820
QY 1482 GACAGAGCTTGAAGAGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 1541
DB 73819 GACAGAGCTTGAAGAGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 73760
QY 1542 GCTTCTGCTCAAAATGTCATGATGATGATGATGATGATGATGATGATGATGATGATGAT 1601
DB 73759 GCTTCTGCTCAAAATGTCATGATGATGATGATGATGATGATGATGATGATGATGATGAT 73700
QY 1602 ACCTGACATGTCAGTTTCAACCATGAGAAAGATTTAGTGTGACATGATGATGATGAT 1661
DB 73699 ACCTGACATGTCAGTTTCAACCATGAGAAAGATTTAGTGTGACATGATGATGATGAT 73640
QY 1662 TCTCTTGAATATTAACAATATCTTCTGTTATCTTGAATGATGATGATGATGATGAT 1721
DB 73639 TCTCTTGAATATTAACAATATCTTCTGTTATCTTGAATGATGATGATGATGATGATGAT 73580

OY	1722	CACTTTTTCGCACTTTTCATATGTAAATPAAAAGAGTGTGTCGCAACAACCTTAATACCTT	1781
Db	73579	CACTTTTTCTGCAATTTTCTTTGTATATPAAAAGAGTTCCTGTACAAAACCTTAAGAC-	73521
OY	1782	TCCTTATACCTGTTTCCCTACCTGATAGTATCAAAAAGAGATTCCTATTAATCTGTCA	1841
Db	73520	TCCTTATACCTTCTTCCCTACCTGATAGTATCAAAAAGAGATTCCTATTAATCTGTCA	73461
OY	1842	GACTATGTTCCCTCGAAATATAGTTCCTTTTATGATCTGAGGGATTAATGACAGTGA	1901
Db	73460	GACTATGTTCCCTCGAAATATAGTTCCTTTTATGATCTGAGGGATTAATGACAGTGA	73401
OY	1902	AGCTCAATTCCTTAATAAGTAGTCTGCTACCTCTAAATTCATATGAATTCAGATATA	1961
Db	73400	AGCTCAATTCCTTAATAAGTAGTCTGCTATCTCAATTCATATGAATTCAGATATA	73341
OY	1962	AAGCAAAATATATGACTTAAAGAGAGATTCCTCCTTCAATPAAAACAGTCTTAAGAAATTG	2021
Db	73340	GAGCAAAATATATGCTTAAAGAGAGACTCTCTCTTCAATPAAAACAGTCTCACTATTTG	73281
OY	2022	TTTTATGAATAGCCCTCCTGCTATTTGTCCACAGCATGAGATAGTGGCCCTTGTT	2081
Db	73280	TTTTATPAAAAGTCTTCCCTGCTATTTGTCCACAGCATGAGATAGTGGCCCTTGTT	73221
OY	2082	TCCTAGTAAAGACATCGTGCCCTTCCCTTGAGAACGTGTAATTCCTATTTAGCTCT	2141
Db	73220	TCCTAGTAAAGACATCGTGCCCTTCCCTTGAGAACCTTTTAATGTCTATTTAGCTCT	73161
OY	2142	TCCTGACCTATATGACTAGTAGAGAGCCTATATATGTCCACCAAGTTTCAATTTGGCC	2201
Db	73160	TCCTGACCTATATGAGACAGTAGAGAGCCCAATAATGTG---CCCAAGTCTATTTGGCC	73104
OY	2202	ATTGGAACCTCAATATTTGAATTTTAAAGTGAATAATCTGGAATAACCTATTAATTTCA	2261
Db	73103	ATTGGAACCTCAATATTTGTTTCAAAAGTGAATAATCTTGAATAACCTATTAATTTTG	73044
OY	2262	CTTACAGATTCCTTTCAGTTGTAGAGAAATTCCTTACCTTCCAGGTTTGTATTAATTGT	2321
Db	73043	CTTACAGACTCTTTCAGTTGTAGAGAAATTCCTTACCTTCCAGGTTTGTATTAATTGT	72988
OY	2322	TCGTATTTTACCTTCACTTCACTTATTTATGCTGTTTACATGAGAACCAACTGAACA	2381
Db	72983	TCGTGCTGTATCTTCACTTCACTTATTTATGAGTATTTGAATACTTGAATACTGAATTA	72924
OY	2382	TCGTGACCTTCCATGACATCTCAATTAATGATCTGATATATTAATTAATCACTGGTACA	2441
Db	72923	TCGTGACCTTTCATGCTATATGCTCAATTAATGATCTTGAATTAATTAATCACTGGTATA	72864
OY	2442	GAAATTCGATATACGTCTGTGACATATACGAACCTGGAATATTTGTCTAAGAAAATAG	2501
Db	72863	GAAATTCGATATACATGCT-TGACATATGATGAACTGTGAACATTCGTATAGAAAATAG	72805
OY	2502	CCAGACGCCAAACATATTTGTAAGTTCAATTCATAGAGTATCCAAATTAGAAATTC	2561
Db	72804	CCAGATATCCAAAGAACCA--GTAAATTTCAAAATTCACCGATATCCAAACTAGAAATTCG	72747
OY	2562	TGAACACAGAAAATTAATTTAGAGAGATCTGTGTCTGA 2600	
Db	72746	TGAACACAGAAAATTAATTTAGAGAGATCTGTGTCTGA 72708	
RESULT 8			
LOCUS	AC139482		
DEFINITION	Homo sapiens chromosome 5 clone RP11-583F24, WORKING DRAFT		
ACCESSION	AC139482		
VERSION	AC139482.1	GI:28201512	
KEYWORDS	HTG; HTGS_PHASE1; HTGS_DRAFT; HTGS_ACTIVEFIN.		
SOURCE	Homo sapiens (human)		
ORGANISM	Bukaryaota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;		

REFERENCE	1 (bases 1 to 165490)
AUTHORS	DOE Joint Genome Institute.
TITLE	Sequencing of Human Chromosome 5
JOURNAL	Unpublished
REFERENCE	2 (bases 1 to 165490)
AUTHORS	DOE Joint Genome Institute.
TITLE	Direct Submision
JOURNAL	Submitted (04-FEB-2003) Production Sequencing Facility, DOE Joint Genome Institute, 2800 Mitchell Drive, Walnut Creek, CA 94598, USA
COMMENT	-----Genome Center

```

Center: Joint Genome Institute
Center Code: JGI
Web site: http://www.jgi.doe.gov
-----
Project Information
Center Project Name: 1518520
Center clone name: RPCR-11_583P2

```

```

-----
Summary Statistics
Consensus quality: 147035 bases at least Q40
Consensus quality: 155210 bases at least Q30
Consensus quality: 158567 bases at least Q20
Estimated insert size: 175000; agarose-fp estimation
Estimated insert size: 164390; sum-of-contigs estimation
Quality coverage: 3.94 in Q20 bases; agarose-fp estimation
Quality coverage: 4.19 in Q20 bases; sum-of-contigs estimation
* NOTE: This is a 'working draft' sequence. It currently
* consists of 12 contigs. The true order of the pieces
* is not known and their order in this sequence record is
* arbitrary. Gaps between the contigs are represented as
* runs of 'N', but the exact sizes of the gaps are unknown.
* This record will be updated with the finished sequence
* as soon as it is available and the accession number will
* be preserved.

```

*	1	1249: contig of 1249 bp in length
*	1250	1349: gap of unknown length
*	1350	3680: contig of 2331 bp in length
*	3681	3780: gap of unknown length
*	3781	6668: contig of 2888 bp in length
*	6669	6768: gap of unknown length
*	6769	10866: contig of 4038 bp in length
*	10867	10966: gap of unknown length
*	10967	17019: contig of 6053 bp in length
*	17020	17119: gap of unknown length
*	17120	22433: contig of 5314 bp in length
*	22434	22533: gap of unknown length
*	22534	34645: contig of 12112 bp in length
*	34646	34745: gap of unknown length
*	34746	49173: contig of 14428 bp in length
*	49174	49273: gap of unknown length
*	49274	67058: contig of 17785 bp in length
*	67059	67158: gap of unknown length
*	67159	84928: contig of 17770 bp in length
*	84929	85028: gap of unknown length
*	85029	114464: contig of 22436 bp in length
*	114465	114564: gap of unknown length
*	114565	163590: contig of 50326 bp in length

```

FEATURES
    source
        location/Qualifiers
            1..165490
                /organism="Homo sapiens"
                /mol_type="Genomic DNA"
                /db_xref="taxon:9606"
                /chromosome="5"
                /clone="RP11-583F24"
                /clone_1fb="RPC1 human BAC library 11"

```

ORIGIN	Query Match	Score	DB 2	Length
Best Local Similarity	88.3%	Pred. No. 0		
Matches 1891; Conservative	0	Mismatches 218;	Indels 33;	Gaps 9
by	466	GGGTCAATCAGACTGGGGTTTCTGAGCAGATGATTCACCAATCCAGTCTTGAGTACAGAAC 525		

Db	121579	GGGACACCAAGACTACCGTTTCTGAGGTGGATTCGAAGCAACCAAGCCTTGGATGCAAGAAC	121638
QY	526	TGAACCAATCAACGAGCGTAGAGAGCTCTTGTCTACAGACAGACCTTGAAGCTTCAAG	585
Db	121639	TGACACCAATTAACAGAACTGAGGAGACTCTTGTCTACAGAGAGACCTTGAAGCCTATGG	121698
QY	586	GGCGTAGCGTGATCGTTTCCCTTGGTCGGGCTGACAGGAAACGGGGTTGGTCTGGGCTCC	645
Db	121699	GGCGTAGCGTGATCAATTCCTTGGTCAAGCTGACAGGAAACGGGGTTGGTCTGGGCTCC	121758
QY	646	TGGCGTGGCGCATGCGCAGAGAAAGCGTGTCTCATCTACATACCTCTCAACTGGTGGCGCG	705
Db	121759	TGGGCTTCCGATGGCGAGAGAACGGCGTCTCATCTACATCTCTCAACTGGTGGCGGAG	121818
QY	706	ACTTCTCTTCTTGAAGCGGCAACATTAAATGTTGGCGTTAGCGCTCATCAATATCCGC	765
Db	121819	ACTTCTCTTCTCTAGCGGCGCAAGTTATCATTCGCGCTCACTCTCATCAATATCTGTC	121878
QY	766	ATCCCATCTCCAAATTCCTCACTCGTGTATGACCTTTCCTACTTTAATAGGCTTAAGCA	825
Db	121879	ATCCCATCTCCAAATTCCTCATTCCTGTATGACCTTTCCTACTTTAATAGGCTTAAGCT	121938
QY	826	TGCTGAGCGGCATCTGACACCGAGCGGTGCTGTCCATCCGTGGGCCATCTGGTACACT	885
Db	121939	TTCTGAGTGCATGAGCACCGAGCGGTGCTGTGCCTCTGTGGCCCATCTGGTACCGCT	121998
QY	886	GCCGC - -GGCCCAATACCTGTCAATCGGTCAATGTGTCTCTGCGCCCTGTCCCT	943
Db	121999	GCTCTCTCCCCCACAACCTGTGACGGGTGCTGTGTCTTGTCTTGGGCCCTGTCCCT	122058
QY	944	GCTGCGAGTATCTGAGTGTGATGTTCTGTGTACTTCTGTTTAAAGTGTGTATCTGT	1003
Db	122059	ACTGCGGAGCAATCTGAGTGTGATGTTCTGTGTACTTCTGTTTAAAGTGTGTATCTAT	122118
QY	1004	TTGGTGTAAACGTGAGATTTGATTACATACGCGTGGGCTGTTTAAATGATGGTCT	1063
Db	122119	TTGGTGTAAACCATGATTTGATTACATACGAGTGGTGGCTGATTTTAAATGATGGTCT	122178
QY	1064	CTGTGGGTCAAGCCTGTGTCTGTGTGACAGAAATCTCTGTGATATCCCGAAGATGCGCT	1123
Db	122179	CTGTGGGTCAAGCCTGTGTCTGTGTGATTTGAAGTATCTCTGTGATATCCCGAAGATGCGCT	122238
QY	1124	GACCAAGCTGATGAGCAATCTCTCTCAACAGTGTGTCTCTCTCTGTGTGAGCCTGCC	1183
Db	122239	GACCGGGGTGTGCTGATGAGCAATCTCTCTCTCAACAGTGTGTCTCTCTCTGTGTGAGCCTGCC	122298
QY	1184	CTTGGGCATTCAGTGGGCGCTGTTTCCAGATATCCACCTGGATTTGGAAGCTTATTTTG	1243
Db	122299	CTTGGGCATTCGTTGGGCGCTGTGTCTAATGGAATCCACCTGGATTTGGAATCATTTTCTG	122358
QY	1244	TCAATGTGATCAGATTTTTCATTTTCTGTGCGCTCTTAAACAGAGTGCACACCCATCAT	1303
Db	122359	TCAATGTGATCAGATTTTTCATTTTCTGTGCGCTCTTAAACAGTGTGCACACCCGTCAT	122418
QY	1304	TTACTTCTTCTGTGGGCTCTTTTAAAGCAGCGCTCAAAATAGGCAGAACCTGAAGCTGTCT	1363
Db	122419	TTACTTCTTCTGTGGGCTCTTTTAAAGCAGCGCTCAAAATAGGCAGAACCTGAAGCTGTCT	122478
QY	1364	CCAGAGGGCTCTGCAAGACACGCTTGAAGTGAATGAAGTGAAGGGTGGCTTCTCAGGA	1423
Db	122479	CCAGAGGGCTCTGCAAGACATGCTGAAG - - -TGAAGGTGAAGGGCGGCTTCTGAAGG	122534
QY	1424	AACCTGGAGCTGTGGGGAAGAGATTTGAGCAGTAGAGGAAGAACCTGTGGCCCT - - -GT	1479
Db	122535	AACCTGGAGCTGTGGGGAAGAGATTTGGGACAGTAGAGGAAGAACCTGTGGCCCTGTGAGT	122599
QY	1480	CAGACAGGACTTTGAAGCAATGTGCGCCTGTCCACCCCTTGAACAATTATGACTTTTCT	1539
Db	122595	CAGACGGGACTTTTGAAGCAATGTGCGCCTGTCCACCCCTTGAACAATTATGACTG - GTTTTCT	122655
QY	1540	TAGCTTCTGTCTCAGAAATGTCTCAGTGTCCCTCAAGGCTTTTCCAAATGATGTTTATC	1599

Db	122554	TAGGTGGCTCAGAAATGTCTCAGGGGTAACTCAAGTGTGTGAAATGATGTTTATC	1227713
Qy	1600	TAACCTGCAACTTGGAGTTTCCACCCATGGAAGAATTAAGTGTGACAGTACAATGTTTTG	1659
Db	122714	TATACTGACAGTTCCAGTTTCTATCCACAGAAAGATTAAGTGTGACAGTACAATGTTTAG	122773
Qy	1660	ATTCTCCGTGAATTAACCATATACATTTTCCGTGTATCTTGCATGMAATCTTCCACTAGT	1719
Db	122774	ATTCTCTGTGATTTACCAACATTTTTCCTGTATCTTAC-----ACTG	122819
Qy	1720	AACACTTTTTCGCACTTTTCATTTGTAATTAAGAAAGAGTTGCTGTCCACAACCTTAACCT	1779
Db	122820	AAACTTTTTCGCAATTTTTCATTTGTAATTAAGAAAGAGTTTCCCTCCACAACCTTAAC---	122876
Qy	1780	CTTCTTATFACTTGTGTTTCCATCCATGATATGATATCAAAAAGGAAGATTCCTATTAATCTGT	1839
Db	122877	CTCATCTTTTAACGTTTCCATCCATGATTTGTATCAAAAAGGAAGTTTCTTATTTGATCTCT	122936
Qy	1840	CAGACTATGTTCCCTGAAATATCATGTTTCCCTTTATATGACTGAGGCAATTACTGCAGTTG	1899
Db	122937	CAGACTATGTTCCCTG-AAAACATATTTCCCTTATATGACTGGAAGCAATTACTGAAGTTG	122995
Qy	1900	GAAGCTCAATCTTATATATAGAGTTCTGTCTACTCTTAATTCATATGAATTCACAGATA	1959
Db	122996	GGAAGTGAATCTGATATAGAGAGTTCTGTTATCTTACATTCACACTGAATTCACAGATA	123055
Qy	1960	TAAAGCAAAATATATGACTTTAGAGAGATTTCTCCCTCATATAAAACAAGCTTTAGAAAT	2019
Db	123056	TAAAGCAAAATATATGCTTCTTAGAGACAGACTTTCTTCATATAAAACACTCTCACAAAT	123115
Qy	2020	GGTTTTATGAAATAGCCCTCCTGTCATTTGTCCACAGATGGTGCATGTTGGCCTGTG	2079
Db	123116	GGTTTTATGAAAGGCTCTCCTGTCATTTGTCCACAGATGGTGCACACTGGCCTGTG	123175
Qy	2080	TTTCTAGTAAAGACATATGTGGCCCTTCCCTTGAAGACTGTAAATCTTATTTAGCT	2139
Db	123176	TTTCTAGTAAAGACACATATGGCCCTTCCCTTGAAGACTGTAAATCTTATTTAGCT	123235
Qy	2140	CTTCTGAGCTATATGACTAGTGAAGGCTATATTAATATGTCCACACAGTTTCATTTTG	2199
Db	123236	CTTCTGAGCTATATATGACACAGTGAAGGACCTGTAAATATGTCCACACAGTTTCATTTTG	123295
Qy	2200	CCATTGGAACCTCAATTTATTTTAAAGGGAATATCTTGAAGAACATTTATTTAT	2259
Db	123296	CCATTGGAATATGCGATATGATTAATTAAGTGAACATATCTTGAAA--AATCATTTAT	123352
Qy	2260	CACCTTACAGATTTCTTCAAGTTGAGGAATCTTCAATCTTCCAGGTTTGTATTAAT	2319
Db	123353	CACCTTACAGATTTCTTCAAGTTGAGGAATCTTCTTATCTTCCAGGTTTGTGACAAAT	123412
Qy	2320	GTTCTGATTTGATCTTTCAGTTAGTTTATGTGCTTTATCATGAGAACCAAACTGAATA	2379
Db	123413	TTTCTGGCTGTAACCTTTCAGTTAGTTTATGTGCTTTCACGTGAGGAACCAAACTGAATA	123472
Qy	2380	CATCTGACCTTCCATGACATCTCAATTTATGTATCTGGAATATTAATTAACAGTTGTGTA	2439
Db	123473	TATCTGACCTTCCATGTCAATATGTCAATTAATGTATCTGGAATATTAATTAACAGTTGTGTA	123532
Qy	2440	CAGAAATCTGATACATGTGCTGACATACATGAACCTGGAATATTTGTGCTAAGGAATA	2499
Db	123533	TAGAAATCTTAATACATGTGCTGACATATGATGAACCTGGAACATTTGTGCTAAGGAATA	123592
Qy	2500	AGCCAGACGCA-AACAATAATTGTATAGTTCAATTTCTATGAGGTATCCAATTTAGGAAT	2558
Db	123593	AGCCAAACCAAGAACATATTTGTATAGTTCAAAATTTTATGAGATATTTGAACCTAGGAAT	123653
Qy	2559	TCTTGAACACAGAAATTAATTTAGAGGATCTGTGTCTGGA	2600
Db	123653	TTGTGAACACAGAAATTAATTTAGAGGAGCTGTGTCTGGA	123694

LOCUS CO815414 2040 bp DNA linear PAT 03-JUN-2004
DEFINITION Sequence 1 from Patent WO2004042402.
ACCESSION CO815414
VERSION CO815414.1 GI:48144069
KEYWORDS
SOURCE Homo sapiens (human)
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Homidae; Homo.
REFERENCE 1
AUTHORS Golz, S., Brueggemeier, U., Geerts, A. and Thiele, R.
TITLE Diagnostics and Therapeutics for diseases associated with human
JOURNAL mas-related gene x1 (mrx1)
Patent: WO 2004042402-A 1 21-MAY-2004;
Bayer Healthcare AG (DB)
FEATURES
Source location/Qualifiers
1.2040
/organism="Homo sapiens"
/mol_type="unassigned DNA"
/db_xref="taxon:9606"
ORIGIN
Query Match 56.7%; Score 1483.8; DB 6; Length 2040;
Best Local Similarity 91.4%; Pred. No. 0;
Matches 1596; Conservative 0; Mismatches 147; Indels 4; Gaps 2;
QY 462 TCATGGGTCATCAGACTGGGGTTTCTGAGCATGATTCACACCATCCAGCTTGGGTACA 521
DB 298 TCCAGGGTCACCAAGCTAGGGTTTCTGAGCATGATTCACACCATCCAGCTTGGGTACA 357
QY 522 GAACCTGACCAATCAACGAGCGTAGAGAGACTCTTGTCTACAGAGACCTGAGCTTC 581
DB 358 GAACCTGACCAATCAACGAGCGTAGAGAGACTCTTGTCTACAGAGACCTGAGCTTC 417
QY 582 ACCGGGCTGAGTGATGTTCCCTTGGGCGCGTAGAGAGAGAGAGAGAGAGAGAGAGAG 641
DB 418 ACCGGGCTGAGTGATGTTCCCTTGGGCGCGTAGAGAGAGAGAGAGAGAGAGAGAGAG 477
QY 642 CTCCTGGGCTGCGGATGCGGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 701
DB 478 CTCCTGGGCTGCGGATGCGGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 537
QY 702 GCCGACTTCTCTTCTTAGCGGCAATTAATATGTTGCGCTTACGCTCATCATATC 761
DB 538 GCAAGACTTCTCTTCTTAGCGGCGGCTTAATATGTTGCGCTTACGCTCATCATATC 597
QY 762 CGCCATCCCATCTCAAAATCTCACTCTGATGATCACTTCCCTACTTTATAGGCTTA 821
DB 598 CGCCATCCCATCTCAAAATCTCACTCTGATGATGATGATGATGATGATGATGATGATG 657
QY 822 AGCATGCTGAGCGCATGAGACGAGCGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 881
DB 658 AGCTTTCTGAGGCGGATGAGACGAGCGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 717
QY 882 CACTGCGCGCGCGGAGATCCTGTCATGATGATGATGATGATGATGATGATGATGATGAT 941
DB 718 CGTGCACACCGCGGAGACGAGCGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 777
QY 942 CTGCTGCGGAGATCTGAG 1001
DB 778 CTGCTGCGGAGATCTGAG 837
QY 1002 GTTGGTGTGAACCTCAGATTTCAATTCAGATGCGCTGCTGCTGCTGCTGCTGCTGCTG 1061
DB 838 GCTTGTGTGAACCTCAGATTTCAATTCAGATGCGCTGCTGCTGCTGCTGCTGCTGCTGCT 897
QY 1062 CTCTGTGTGTGAGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 1121
DB 898 CTCTGTGTGTGAGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 957
QY 1122 CTGACAGAGCTGTAGTACCATCTCTCTCAAGAGCTGTCTCTCTCTCTCTCTCTCTCTG 1181
DB 958 CTGACAGAGCTGTAGTACCATCTCTCTCAAGAGCTGTCTCTCTCTCTCTCTCTCTCTG 1017

QY 1182 CCGTTTGGCATTCAGTGGGCGCTGTTTCCAGATTCACCTGATTTGGAAGCTTATTT 1241
DB 1018 CCGTTTGGCATTCAGTTTTCTTATTTTATGATTCACGAGGAGGAGGCTTATTT 1077
QY 1242 TGTGATGTGATCTGATTTTCCATTTTCTGTCCTGCTTAAACAGAGTCCCAACCCATC 1301
DB 1078 TGTGATGTGATCTGATTTTCCATTTTCTGTCCTGCTTAAACAGAGTCCCAACCCATC 1137
QY 1302 ATTACTCTTGTGGGCTGCTTGTAGGAGGCTCAAAATAGGAGAACTGAACTGAT 1361
DB 1138 ATTACTCTTGTGGGCTGCTTGTAGGAGGCTCAAAATAGGAGAACTGAACTGAT 1197
QY 1362 CTCGAGAGGCTCTGACAGAGACGCTGAGGTGATGAGAGAGAGAGAGAGAGAGAGAG 1421
DB 1198 CTCGAGAGGCTCTGACAGAGACGCTGAGGTGATGAGAGAGAGAGAGAGAGAGAGAG 1257
QY 1422 GAAACCTGAGAGCTGTGGGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 1481
DB 1258 GAAATCTGAGAGCTGTGGGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 1317
QY 1482 GACAGAGCTTTGAGAGAGATGCTGCTGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 1541
DB 1318 GACAGAGCTTTGAGAGAGATGCTGCTGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 1377
QY 1542 GCGTTTGTGCTGAGAAATGCTCAAGTGTCTCAAGGCTTTCGATATGATGTTATCTA 1601
DB 1378 GCGTTTGTGCTGAGAAATGCTCAAGTGTCTCAAGGCTTTCGATATGATGTTATCTA 1437
QY 1602 ACCGACAGTTCAGATTTTCAACCAATGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 1661
DB 1438 ACCGACAGTTCAGATTTTCAACCAATGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 1497
QY 1662 TCTCCTGATATTAACAATATCTTCCCTGATATCTTGCAGTGAATCTTCTCACTGAA 1721
DB 1498 TCTCCTGATATTAACAATATCTTCCCTGATATCTTGCAGTGAATCTTCTCACTGAA 1557
QY 1722 CACTTTTCTGCACTTTTCAATGATTAATAAGAGAGTGTCTGCAAAACCTTAAACTCT 1781
DB 1558 CACTTTTCTGCACTTTTCAATGATTAATAAGAGAGTGTCTGCAAAACCTTAAACTCT 1616
QY 1782 TCTTTATATCTGTTTCCATCAGATGATGATGATGATGATGATGATGATGATGATG 1841
DB 1617 TCTTTATATCTTCTTCACTGATGATGATGATGATGATGATGATGATGATGATGAT 1676
QY 1842 GACTATGTTCCCGTGAATATCATGTTCCCTTTATGACTGAGAGCATTAAGAGTTGA 1901
DB 1677 GACTATGTTCCCGTGAATATCATGTTCCCTTTATGACTGAGAGCATTAAGAGTTGA 1736
QY 1902 AGCTCAATTTCTTAATAGTATGATGATGATGATGATGATGATGATGATGATGAT 1961
DB 1737 AGCTCAATTTCTTAATAGTATGATGATGATGATGATGATGATGATGATGATGAT 1796
QY 1962 AAGCAAAATATGACTTATGAGAGAGATTTCCCTTCAATAAAACAGCTTAAGAA 2021
DB 1797 AAGCAAAATATGACTTATGAGAGAGATTTCCCTTCAATAAAACAGCTTAAGAA 1856
QY 2022 TTTTATGAATAGCCCTCTCTCTGCTCAATTTGTCACAGATGATGATGATGATGAT 2081
DB 1857 TTTTATGAATAGCCCTCTCTCTGCTCAATTTGTCACAGATGATGATGATGATGAT 1916
QY 2082 TCTATGAATAGCAATGATGATGATGATGATGATGATGATGATGATGATGATGAT 2141
DB 1917 TCTATGAATAGCAATGATGATGATGATGATGATGATGATGATGATGATGATGAT 1976
QY 2142 TCTTGAATATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 2201
DB 1977 TCTTGAATATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 2033
QY 2202 ATTGGA 2208
DB 2034 ATTGGA 2040

RESULT 10
LOCUS AX299143 2040 bp DNA linear PAT 26-NOV-2001
DEFINITION Sequence 15 from Patent WO0183555.
ACCESSION AX299143
VERSION AX299143.1 GI:17129051
KEYWORDS
SOURCE Homo sapiens (human)
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.
REFERENCE
1 Anderson, D.J., Dong, X., Zylka, M., Han, S.K. and Simon, M.
Pain signaling molecules
Patent: WO 0183555-A 15 08-NOV-2001,
CALIFORNIA INSTITUTE OF TECHNOLOGY (US)
FEATURES
source
1. .2040
/organism="Homo sapiens"
/mol_type="unassigned DNA"
/db_xref="taxon:9606"
328..1296
/note="unassigned protein product"
/codon_start=1
/protein_id="CAD12706.1"
/db_xref="GI:17129052"
/translation="MDPTISTLDLTPINGTEBTLCKYKQTLSTLTLCIVSLVGLTG
NAVAVLMLGGRMRNAPSIYILNLAAADPLFLSGRLVLSLPSISPHITSKILVPM
MPSYFAGLSPLSAVTERCLSLVMPWIRCHPRLHSAVVCVLSLMLSLRLIENWL
CGFLFSGADAMCQSDPFIYAVMLIFLCVLCGSLVLLRLICGSSKRIPLRLYVLI
LTLVLFVLCGPRGIOPFLVIMHVRVLPCHVLSVPLSLNLSNAPDIYFVFG
SPFRQRQRQVLMKLVLRALODASVDEGGGLPERILIELSGSRLEQ"

ORIGIN

Query Match 56.7%; Score 1483.8; DB 6; Length 2040;
Best Local Similarity 91.4%; Pred. No. 0;
Matches 1596; Conservative 0; Mismatches 147; Indels 4; Gaps 2;

QY 462 TCATGGGTCAATCAGATCGGGTTTCTGAGCATGATTCACCATCCAGTTGGGATCA 521
DB 298 TCCAGGGGTCAACAGATCAGGGTTTCTGAGCATGATTCACCATCCAGTTGGGATCA 357
QY 522 GAATGACACCAATCAACGAGCGTGAAGAGATCTCTGATCAAGACGAGCCCTGAGCTC 581
DB 358 GAATGACACCAATCAACGAGCGTGAAGAGATCTCTGATCAAGACGAGCCCTGAGCTC 417
QY 582 ACCGGGCTGACATGATCGTTCCCTTGTGCGCTGACAGAGAAACGGGTTGCTCTGG 641
DB 418 ACCGGGCTGACATGATCGTTCCCTTGTGCGCTGACAGAGAAACGGGTTGCTCTGG 477
QY 642 CTCCTGGGCTGACGAGATCGGAGAAACGCTGTCTTCATCTCAATCTTCACCTGCTGCG 701
DB 478 CTCCTGGGCTGACGAGATCGGAGAAACGCTGTCTTCATCTCAATCTTCACCTGCTGCG 537
QY 702 GCGGACTTCTCTTCTTAAGGCGCAATATATGTTGCGCGTAAAGCCTCATGAATATC 761
DB 538 GCGGACTTCTCTTCTTAAGGCGCGCCTTAATATATTCCTGTAAAGCTCATGAATATC 597
QY 762 GCGCATCCCATCTTCAAAATCTCAGTCTGTGATGACCTTCCCTATTTATAGGCTTA 821
DB 598 CCCCATACATCTTAATATCTTATCTCTGTATGATGATTTTCTATTTGACAGGCTG 657
QY 822 AGCATGTGAGCGCATCAGACCGAGCGCTGCTGTTCATCTGTGGCCATCTGTGTAC 881
DB 658 AGCTTTTGTAGTGTGCGTGAAGACCGAGCGCTGCTGTTCCTGTGTGGCCATCTGTGTAC 717
QY 882 CACTGCGCGCGCCCAATACCTGTCAATCGGTCAATGATGATGCTGTGGGCGCTGTGC 941
DB 718 CACTGCGCGCGCCCAATACCTGTCAATCGGTGATGATGCTGTGGGCGCTGTGC 777
QY 942 CTGCTGCGAGATATCTGAGATGATGTTCTGTGACTTCTGTGATGATGATGATCT 1001

DB 778 CTGCTGCGAGATATCTGAGATGATGATGATGATGATGATGATGATGATGATGATCT 837
QY 1002 GTTGGGTGAAACGTGATTTCAATACATCGCGGCTGTTTATGATGATGATGATGATGAT 1061
DB 838 GCTTGGGTGAAACGTGATTTCAATACATCGCGGCTGTTTATGATGATGATGATGATGAT 897
QY 1062 CTCTGTGGGTGAAACGTGATTTCAATACATCGCGGCTGTTTATGATGATGATGATGATGAT 1121
DB 898 CTCTGTGGGTGAAACGTGATTTCAATACATCGCGGCTGTTTATGATGATGATGATGATGAT 957
QY 1122 CTGACACGCTGTGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 1181
DB 958 CTGACACGCTGTGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 1017
QY 1182 CCTTGGGCAATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 1241
DB 1018 CCTTGGGCAATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 1077
QY 1242 TGTGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 1301
DB 1078 TGTGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 1137
QY 1302 ATTTACTTCTTCTGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 1361
DB 1138 ATTTACTTCTTCTGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 1197
QY 1362 CTCGAGGCGCTGCGAGACGCGCTGAGTGTGATGATGATGATGATGATGATGATGATGAT 1421
DB 1198 CTCGAGGCGCTGCGAGACGCGCTGAGTGTGATGATGATGATGATGATGATGATGATGAT 1257
QY 1422 GAAACCTGAGAGCTGTGCGGAGACGATGATGATGATGATGATGATGATGATGATGATGAT 1481
DB 1258 GAAACCTGAGAGCTGTGCGGAGACGATGATGATGATGATGATGATGATGATGATGATGAT 1317
QY 1482 GACAGACCTTGAAGACATGCTGCTGCAACCTTGAACATTAATGATTTTCTTA 1541
DB 1318 GACAGACCTTGAAGACATGCTGCTGCAACCTTGAACATTAATGATTTTCTTA 1377
QY 1542 GCTTTCGCTGCAAAATGCTGATGATGATGATGATGATGATGATGATGATGATGATGAT 1601
DB 1378 GCTTTCGCTGCAAAATGCTGATGATGATGATGATGATGATGATGATGATGATGATGAT 1437
QY 1602 ACCGTGACATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 1661
DB 1438 ACCGTGACATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 1497
QY 1662 TCTCTGATATTAACAATATCTTCTGATGATGATGATGATGATGATGATGATGATGATGAT 1721
DB 1498 TCTCTGATATTAACAATATCTTCTGATGATGATGATGATGATGATGATGATGATGATGAT 1557
QY 1722 CACTTTTCTGCAATTTCTTGAATAAAGAGTGTGCTGCAAAACCTTAACACTCT 1781
DB 1558 CACTTTTCTGCAATTTCTTGAATAAAGAGTGTGCTGCAAAACCTTAACACTCT 1616
QY 1782 TCTTTATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 1841
DB 1617 TCTTTATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 1676
QY 1842 GACTATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 1901
DB 1677 GACTATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 1736
QY 1902 AGCTGATTTCTTAATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 1961
DB 1737 AGCTGATTTCTTAATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 1796
QY 1962 AAGCAATATATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 2021
DB 1797 GAGCAATATATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 1856
QY 2022 TTTTATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 2081
DB 1857 TTTTATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 1916

QY 2082 TCTAGTAAAGCAATCGTGGCCCTTCCCTTGAGAGACTGTGTAAGTTCTTATTACTCT 2141
DB 1917 TCTAGTAAAGCAATCGTGGCCCTTCCCTTGAGAGACTGTGTAAGTTCTTATTACTCT 1976
QY 2142 TCTGAGACTAATGAACTAGTGGAGAGCCCTAATAATGTCCACCGCTTCAATTTGGCC 2201
DB 1977 TCTGAGACTAATGAACTAGTGGAGAGCCCTAATAATGTCCACCGCTTCAATTTGGCC 2033
QY 2202 ATTGGA 2208
DB 2034 ATTGGA 2040

RESULT 11
AY042213 2040 bp DNA linear pri 14-SRP-2001
LOCUS Homo sapiens G protein-coupled receptor (MRGX1) gene, complete cds.
ACCESSION AY042213
VERSION AY042213.1 GI:15546061
KEYWORDS
SOURCE Homo sapiens (human)
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE 1 (bases 1 to 2040)
AUTHORS Dong, X., Han, S., Zylka, M.J., Simon, M.I. and Anderson, D.J.
TITLE A diverse family of Gpcrs expressed in specific subsets of
neocortical sensory neurons
Cell 106 (5), 619-632 (2001)
JOURNAL MEDLINE 21435808
PUBMED 11551509
REFERENCE 2 (bases 1 to 2040)
AUTHORS Dong, X., Han, S.-K., Zylka, M.J., Simon, M.I. and Anderson, D.J.
TITLE Direct Submission
Submitted (19-JUN-2001) Division of Biology, 216-76, California
Institute of Technology, 1201 E. California Blvd., Pasadena, CA
91125, USA

FEATURES
source location/Qualifiers
1..2040
/organism="Homo sapiens"
/mol_type="genomic DNA"
/db_xref="taxon:9606"
<328..>1296
/gene="MRGX1"
<328..>1296
/gene="MRGX1"
/product="G protein-coupled receptor"
328..1296
/gene="MRGX1"
/note="Mas-related gene X1"
/codon_start=1
/product="G protein-coupled receptor"
/protein_id="AAK91804.1"
/db_xref="GI:15546062"
/translation="MDPTISTDTEITPINTERTLCYKQTLSTVLTCTIVAGLTG
NAVYMLIGCKMRNRSFIYIINLAADPLFSGRLYSLSFISIPHSIKILYPM
MESYFAGSPLSAVSTERCLSVLMPIMWRCHRPHTLSAVCVLMAISLRILSLEML
CGPLSGADSAMCOTSDPTVAMLIPLCVLGGSSIVILIRILGSRKPLRLVYTI
LITVIVPLCGIPFGIOTFFPLMTHYDREVLCHVHLVIFLSALNSANPILITFFVG
SFRORONKMLVLRALQDASEVDEGGQLPEBILISGRLEQ"

ORIGIN
Query Match 56.7%; Score 1483.8; DB 9; Length 2040;
Best Local Similarity 91.4%; Pred. No. 0;
Matches 1596; Conservative 0; Mismatches 147; Indels 4; Gaps 2;

QY 462 TCATGGGTCATCAGACTGGGCTTTCTGAGCATGATCAACATCCAGCTTGGGTACA 521
DB 298 TCACGGGTCACAGACTAGGCTTTCTGAGCATGATCAACATCCAGCTTGGGTACA 357
QY 522 GAAGTACCAATCAAGGAGCTGAGAGACTCTTGTCTACAGAGACCTGAGCTTC 581

DB 358 GAAGTACCAATCAAGGAGCTGAGAGACTCTTGTCTACAGAGACCTGAGCTTC 417
QY 582 ACAGGAGTGAAGCTGATGCTTTCCCTTGCGGCTGACAGAGAAACGCGTGTGCTG 641
DB 418 ACAGGAGTGAAGCTGATGCTTTCCCTTGCGGCTGACAGAGAAACGCGTGTGCTG 477
QY 642 CTCCTGGGCTGCGCATGGCGAGAAAGCTGTCTTCATCTACATCTCAACCTGGTGG 701
DB 478 CTCCTGGGCTGCGCATGGCGAGAAAGCTGTCTTCATCTACATCTCAACCTGGTGG 537
QY 702 GCGGACTTCCTTCTTCTTGAAGGCGCAATTAATGTCGCGGTTACGCTCATCAATTC 761
DB 538 GCGGACTTCCTTCTTCTTGAAGGCGCGCTTATATATCTCTGTATAGCTTATCAATTC 597
QY 762 GCGCATCCCATCTCCAAATCTCTGATGATGATGATGATGATGATGATGATGATGATG 821
DB 598 GCGCATCCCATCTCCAAATCTCTGATGATGATGATGATGATGATGATGATGATGATG 657
QY 822 AGCATGCTGAGCGCATGACGACCGAGGCTGCTGCTCATCTCTGAGCCCATCTGTTAC 881
DB 658 AGCATGCTGAGCGCATGAGCACCGAGGCTGCTGCTCATCTCTGAGCCCATCTGTTAC 717
QY 882 CACTGCGCGCGCCGAGATACCTGTATGATGATGATGATGATGATGATGATGATGATG 941
DB 718 CACTGCGCACCGCCGAGATACCTGTATGATGATGATGATGATGATGATGATGATGATG 777
QY 942 CTGCTGCGGAGATCTGAGAGTGAATGATGATGATGATGATGATGATGATGATGATG 1001
DB 778 CTGCTGCGGAGATCTGAGAGTGAATGATGATGATGATGATGATGATGATGATGATG 837
QY 1002 GTTGTGTGTAACGTCAAGATTTCAATCAATCGCGTGTGTTTATATGTGTGTT 1061
DB 838 GTTGTGTGTAACGTCAAGATTTCAATCAATCGCGTGTGTTTATATGTGTGTT 897
QY 1062 CTGTGCGTCCAGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 1121
DB 898 CTGTGCGTCCAGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 957
QY 1122 CTGACAGGCTGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATG 1181
DB 958 CTGACAGGCTGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATG 1017
QY 1182 CCGTTTGGCATCACTGGGCTGTTTTCACAGATCACTGATGATGATGATGATGATGAT 1241
DB 1018 CCGTTTGGCATCACTGGTTCCTATTTTATGATGATGATGATGATGATGATGATGAT 1077
QY 1242 TGTCAATGATCTATGATGATGATGATGATGATGATGATGATGATGATGATGATGATG 1301
DB 1078 TGTCAATGATCTATGATGATGATGATGATGATGATGATGATGATGATGATGATGATG 1137
QY 1302 ATTACTCTTCTGAGGCTCCTTTAGGAGGCTCAAAATAGGAGAACCTGAGCTGTT 1361
DB 1138 ATTACTCTTCTGAGGCTCCTTTAGGAGGCTCAAAATAGGAGAACCTGAGCTGTT 1197
QY 1362 CTCAGAGGCTCTGAGAGACAGCTTGAAGTGAAGTGAAGTGAAGTGAAGTGAAGTGAAG 1421
DB 1198 CTCAGAGGCTCTGAGAGACAGCTTGAAGTGAAGTGAAGTGAAGTGAAGTGAAGTGAAG 1257
QY 1422 GAAACCTGAGAGCTGCGGAGAGAGATGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 1481
DB 1258 GAAACCTGAGAGCTGCGGAGAGAGATGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 1317
QY 1482 GACAGAGCTTGAAGAGATGCTGCTGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 1541
DB 1318 GACAGAGCTTGAAGAGATGCTGCTGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 1377
QY 1542 GCGTTCTGCTCAAGATGCTCAAGTGTCTCAAGTGTCTCAAGTGTCTCAAGTGTCTCA 1601
DB 1378 GCGTTCTGCTCAAGATGCTCAAGTGTCTCAAGTGTCTCAAGTGTCTCAAGTGTCTCA 1437
QY 1602 ACCTGACAGTTCAGATTTCAACCCATGGAAGAGATTAATCTGACAGTCAATGTTGAT 1661
DB 1438 ACCTGACAGTTCAGATTTCAACCCATGGAAGAGATTAATCTGACAGTCAATGTTGAT 1497


```

QY 1662 TCTCCTGATTAATCAATACATTTTCCCTGTAATCTTGACATGAATCTTCTTACTGAA 1721
DB 1498 TCTCCTGATTAATCAATACATTTTCCCTGTAATCTTGACATGAATCTTCTTACTGAA 1557
QY 1722 CACTTTTCTGACATTTTCTGTAATTAATTAAGAGTTGCTGTCACAACTTAAACTCT 1781
DB 1558 CACTTTTCTGACATTTTCTGTAATTAATTAAGAGTTGCTGTCACAACTTAAACTCT 1616
QY 1782 TCTTTTAATCTGTTTCTTCTGTAATTAATTAAGAGTTGCTGTCACAACTTAAACTCT 1841
DB 1617 TCTTTTAATCTGTTTCTTCTGTAATTAATTAAGAGTTGCTGTCACAACTTAAACTCT 1676
QY 1842 GACTATGTTCCCTGTAATTAATTAAGAGTTGCTGTCACAACTTAAACTCT 1901
DB 1677 GACTATGTTCCCTGTAATTAATTAAGAGTTGCTGTCACAACTTAAACTCT 1736
QY 1902 AGCTCAATTTCTTAATTAAGAGTTGCTGTCACAACTTAAACTCT 1961
DB 1737 AGCTCAATTTCTTAATTAAGAGTTGCTGTCACAACTTAAACTCT 1796
QY 1962 AAGCAAAATTAATTAAGAGTTGCTGTCACAACTTAAACTCT 2021
DB 1797 AAGCAAAATTAATTAAGAGTTGCTGTCACAACTTAAACTCT 1856
QY 2022 TTTTAATTAATTAAGAGTTGCTGTCACAACTTAAACTCT 2081
DB 1857 TTTTAATTAATTAAGAGTTGCTGTCACAACTTAAACTCT 1916
QY 2082 TCTAGTAAGAGTAATTAAGAGTTGCTGTCACAACTTAAACTCT 2141
DB 1917 TCTAGTAAGAGTAATTAAGAGTTGCTGTCACAACTTAAACTCT 1976
QY 2142 TCTAGTAAGAGTAATTAAGAGTTGCTGTCACAACTTAAACTCT 2201
DB 1977 TCTAGTAAGAGTAATTAAGAGTTGCTGTCACAACTTAAACTCT 2033
QY 2202 ATTTGAA 2208
DB 2034 ATTTGAA 2040

```

RESULT 12
BC067292
LOCUS BC067292 1697 bp mRNA linear PRI 30-JUN-2004
DEFINITION Homo sapiens G protein-coupled receptor MRCX3, mRNA (cDNA clone
MGC:75450 IMAGE:30383708), complete cds.
ACCESSION BC067292
VERSION BC067292.1 GI:45501076
KEYWORDS MGC.
SOURCE Homo sapiens (human)
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.
1 (bases 1 to 1697)
Strausberg, R.L., Feingold, B.A., Grouse, L.H., Derge, J.G.,
Klausner, R.D., Collins, P.S., Wagner, L., Shennan, C.M., Schuler, G.D.,
Altschul, S.F., Zeeberg, B., Buetow, K.H., Schaefer, C.F., Hsieh, F.,
Hopkins, R.F., Jordan, H., Moore, T., Max, S.I., Wang, J., Hsieh, F.,
Diatchenko, L., Marusina, K., Farmer, A.A., Rubin, G.M., Hong, L.,
Scapellato, M., Soares, M.B., Bonaldo, M.F., Casavant, T.L.,
Schaefer, T.E., Brownstein, M.J., Usdin, T.B., Toehly, S.,
Carninci, P., Prange, C., Raha, S.S., Loquellano, N.A., Peters, G.J.,
Abramson, R.D., Mullaly, S.J., Bosak, S.A., McEwan, P.J.,
Mokerman, K.J., Malek, J.A., Gunaratne, P.H., Richards, S.,
Worley, K.C., Hale, S., Garcia, A.M., Gay, L.J., Hulyk, S.W.,
Villalón, D.K., Muzny, D.M., Sodergren, E.J., Lu, X., Gibbs, R.A.,
Fahey, J., Helton, E., Kettman, M., Madan, A., Rodriguez, S.,
Sanchez, A., Whiting, M., Madan, A., Young, A.C., Shevchenko, Y.,
Bouffard, G.G., Blakesley, R.W., Touchman, J.W., Green, E.D.,
Dickson, M.C., Rodriguez, A.C., Grimwood, J., Schmutz, J., Myers, R.M.,
Butterfield, Y.S., Krzywinski, M.I., Skalek, U., Smalins, D.B.,
Schnerch, A., Schein, J.E., Jones, S.J. and Marra, M.A.

TITLE
JOURNAL
PUBMED
AUTHORS
TITLE
JOURNAL

Generation and initial analysis of more than 15,000 full-length human and mouse cDNA sequences
Proc. Natl. Acad. Sci. U.S.A. 99 (26), 16899-16903 (2002)
12477932
2 (bases 1 to 1697)
Strausberg, R.
Direct Submission
Submitted (12-MAR-2004) National Institutes of Health, Mammalian Gene Collection (MGC), Cancer Genomics Office, National Cancer Institute, 31 Center Drive, Room 11A03, Bethesda, MD 20892-2590, USA
NIH-MGC Project URL: <http://mgc.nci.nih.gov>
Contact: MGC help desk
Email: cgabs-remail.nih.gov
Tissue Procurement: Dr. Michael Brownstein
cDNA Library Preparation: Invitrogen Corp
cDNA Library Arrayed by: The I.M.A.G.E. Consortium (LNL)
DNA Sequencing by: Genome Sequence Centre,
BC Cancer Agency, Vancouver, BC, Canada
info@cgsc.bc.ca
Steve Jones, Sarah Barber, Mabel Brown-John, Yaron Butterfield,
Andy Chan, Steve S. Chand, William Chow, Allison Cloutier, Ruth
Featherstone, Malachi Griffith, Odi Griffith, Ran Guin, Nancy Liao,
Kim MacDonald, Amara Masson, Mike R. Mayo, Josh Moran, Ryan Morin,
Teika Olson, Diana Palmquist, Anca Petrescu, Anna Liisa Prabhu,
Parvaneh Saeedi, JR Santos, Angelique Schnerch, Ursula Skalek,
Duane Smalins, Jeff Stott, Miranda Tsai, George Yang, Jacqui
Schein, Asim Siddiqui, Rob Holt, Marco Marra.

FEATURES
source
1. 1697
/organism="Homo sapiens"
/mol_type="mRNA"
/db_xref="taxon:9606"
/clone="MGC:75450 IMAGE:30383708"
/tissue_type="Testis"
/clone_id="N1H MGC_180"
/lab_host="DH10B"
/note="vector: pCMV-Sport6.1"
1. 1697
/gene="MRCX3"
/db_xref="LOCUSID:117195"
/db_xref="MIM:607229"
362..1330
/gene="MRCX3"
/codon_start=1
/product="G protein-coupled receptor MRCX3"
/protein_id="AAH67292.1"
/db_xref="GI:45501077"
/db_xref="LOCUSID:117195"
/db_xref="MIM:607229"
/translation="MDSITPVLTGTELPINGRETPCYKOTLSGTGTCIVSYALNG
NAVVWLLGGRMRNNAVSIYILNVADFLPSGHIIRSPRLINHPISKILSPWA
TFPFLISMLSIATERCSILMPWYHCRPRYSWCVLMLALSLRISLSEMF
CDFPLSGADSWECSDPFIAMLVFCLVCLVLAIVLILGSRKMPLETRLYVI
LITLVFLGLGPGIOMALFSLRHLDMKVLFCVHIVSVIFLSLNSANDIIVFVG
SFROROROKIVLQALADTPREVDSGGQLPQETLELSGRLEQ"

ORIGIN
Query Match 49.7%; Score 1300.2; DB 9; Length 1697;
Best Local Similarity 99.0%; Pred. No. 3.9e-309;
Matches 1308; Conservative 0; Mismatches 13; Indels 0; Gaps 0;
QY 457 AATCTCATGGGTATGACACTGGGGTTTCTGACATGATTCACCATCCCACTTTGG 516
DB 327 AACCATTAAGGGGTATGACACTGGGGTTTCTGACATGATTCACCATCCCACTTTGG 386
QY 517 GTACAGAACTGACACCAATCAAGGAGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 576

roman,j., Rosetti, M., Roy, A., Santos, R., Schauer, S., Schupack, R.,
Seaman, S., Severy, P., Spencer, B., Strange-Thomann, N., Stojanovic, N.,
Strassman, N., Subramanian, A., Talamas, J., Teffaye, S., Theodore, J.,
Topham, K., Travers, M., Travis, N., Trigilio, J., Vesiliev, H.,
Viel, R., Vo, A., Wilson, B., Wu, X., Wyman, D., Ye, W. J., Young, G.,
Zainoun, J., Zembek, L., Zimmer, A. and Zody, M.

TITLE
JOURNAL
Submitted (26-APR-2002) Whitehead Institute/MIT Center for Genome
Research, 320 Charles Street, Cambridge, MA 02141, USA
4 (bases 1 to 189230)

REFERENCE
AUTHORS
Birren, B., Linton, L., Nusbaum, C., Lander, E., Ali, A., Allen, N.,
Anderson, S., Barta, N., Bastien, V., Bloom, T., Boguslavsky, L.,
Bouckgeater, B., Brown, A., Camarata, J., Campolongo, A., Chang, J.,
Chazaro, B., Choquet, Y., Collangelo, M., Collins, S., Collamore, A.,
Cook, A., Cooke, P., Deatellano, K., Dewar, K., Diaz, J. S., Dodge, S.,
Faro, S., Ferreira, P., Fitzhugh, W., Gage, D., Galagan, J., Gardyna, S.,
Grinde, S., Gord, S., Goyette, M., Graham, L., Grand-Pierre, N.,
Hagos, B., Horton, L., Hulme, W., Iliev, I., Johnson, R., Jones, C.,
Kamat, A., Karatas, A., Kells, C., LaRoque, K., Lamazares, R.,
Landers, T., Lehoczy, J., Levine, R., Lindblad-Toh, K., Liu, G.,
Maclean, C., MacDonald, P., Major, J., Margulis, N., Matthews, C.,
McCarthy, M., McEwan, P., McKernan, K., Meldrum, J., Meneses, L.,
Mhova, T., Mlenga, V., Murphy, T., Naylor, J., Nguyen, C., Nicol, R.,
Norbu, C., Norman, C. H., O'Connor, T., O'Donnell, P., O'Neil, D.,
Oliver, J., Peterson, K., Phunphang, P., Pierre, N., Pollara, V.,
Raymond, C., Retta, R., Rieback, M., Riley, R., Rise, C., Rogov, P.,
Roman, J., Rosetti, M., Roy, A., Santos, R., Schauer, S., Schupack, R.,
Seaman, S., Severy, P., Spencer, B., Strange-Thomann, N., Stojanovic, N.,
Strassman, N., Subramanian, A., Talamas, J., Teffaye, S., Theodore, J.,
Topham, K., Travers, M., Travis, N., Trigilio, J., Vesiliev, H.,
Viel, R., Vo, A., Wilson, B., Wu, X., Wyman, D., Ye, W. J., Young, G.,
Zainoun, J., Zembek, L., Zimmer, A. and Zody, M.

Direct Submission
Submitted (13-MAY-2002) Whitehead Institute/MIT Center for Genome
Research, 320 Charles Street, Cambridge, MA 02141, USA
On May 13, 2002 this sequence version replaced g1:20330974.
All repeats were identified using RepeatMasker:
Smit, A. F. A. & Green, P. (1996-1997)
<http://ftp.genome.washington.edu/RM/RepeatMasker.html>

Genome Center
Center: Whitehead Institute/ MIT Center for Genome Research
Center code: WIRB
Web site: <http://www-seq.wi.mit.edu>
Contact: sequence_submissions@genome.wi.mit.edu

Project Information
Center project name: L21856
Center clone name: 1081_L_13

FEATURES
source

location/Qualifiers
1. 189230
/organism="Homo sapiens"
/mol_type="genomic DNA"
/db_xref="taxon:9606"
/chromosome="11"
/map="11"
/clone="RP11-1081L13"
/clone_lib="RP11 Human Male BAC"
198. 223
/rpt_family="CTCA)n"
1203. 1215
/note="<30 qual SINGL region"
2295. 2378
/rpt_family="L2"
complement(2381. 3394)
/rpt_family="L1MD2"
3396. 4062
/rpt_family="L2"
4172. 4245
/rpt_family="AT_rich"
4411. 4465
/rpt_family="(TA)n"
complement(4482. 4546)
/rpt_family="L1MC4a"

repeat_region complement(4547. 4857)
/rpt_family="AluJo"
complement(4858. 5400)
/rpt_family="L1MC4a"
6131. 6239
/rpt_family="L2"
complement(6654. 7008)
/rpt_family="MLT1B"
7551. 7749
/rpt_family="L2"
8025. 8057
/rpt_family="AT_rich"
9218. 9800
/rpt_family="L1PB1"
9883. 9955
/rpt_family="L1PB1"
complement(11450. 12969)
/rpt_family="L1PA4"
complement(13902. 14364)
/rpt_family="L1TR33"
complement(14660. 15009)
/rpt_family="MER2"
complement(15115. 15369)
/rpt_family="MLT1J"
15517. 15603
/rpt_family="MIR3"
complement(15612. 15692)
/rpt_family="L2"
16230. 16337
/rpt_family="MIR"
complement(16959. 17034)
/rpt_family="L2"
17340. 17464
/rpt_family="MIR"
19511. 19704
/rpt_family="L1MB5"
19705. 19813
/rpt_family="L1MB4"
20596. 20683
/rpt_family="L2"
complement(20686. 20691)
/note="<30 qual SINGL region"
complement(20691. 20874)
/rpt_family="L1MD3"
21261. 21416
/rpt_family="MIR"
complement(22006. 22307)
/rpt_family="L1P4"
complement(22308. 22559)
/rpt_family="ALUSg1"
complement(22560. 24537)
/rpt_family="L1P4"
24532. 24718
/rpt_family="L1"
complement(24717. 25049)
/rpt_family="L1P4"
complement(25039. 25908)
/rpt_family="L1PA4"
complement(25885. 27145)
/rpt_family="L1"
complement(27179. 29623)
/rpt_family="L1P4"
complement(29632. 30526)
/rpt_family="L1"
complement(30530. 31245)
/rpt_family="L1MBC"
complement(31246. 31555)
/rpt_family="ALUSx"
complement(31556. 31800)
/rpt_family="L1MBC"
complement(31853. 32053)
Query Match 46.9%; Score 1228.2; DB 9; Length 189230;

Best Local Similarity 78.7%; Pred. No. 5,1e-291;
Matches 1695; Conservative 0; Mismatches 403; Indels 55; Gaps 17;

```
QY 466 GGGGATAGAGACTGGGGTTTCTGACATGATGATTAACCATCCAGCTTGGGTAACGAGC 525
Db 17535 GGGGACACGACTAGGGTTTCTGTGATGATCAACCATCCAGCTTGGGTAACGAGC 17594
QY 526 TGACACCAATCAACGAGAGTGAAGACTCT-----TGCTAACAGACAGCCCTGA 576
Db 17595 TGACATCATATGATGAGACCGAGAGACCCATCTTCAACATGTGGCAATGAGATCATGA 17654
QY 577 GCTTACGAGGCTGACGTGATGCTTTCTCTGTGCGCTGACAGAGAAACGGGT--TGT 634
Db 17655 TCTCATGTGCTGCTCTCATGCTGATGACCTGTGACAGCTGAGAGAAATGAGTCAATT 17714
QY 635 GCTGCTGCTGCGGCTGCGGATGCGAGAGAGCGTCTCATCTAATCATCTCAACCT 694
Db 17715 CCTGCTGCTGCGGATGCGGATGCGAGAGAGCGTCTCATCTAATCATCTCAACCT 17774
QY 695 GGTGCGGCGGCACTTCTCTTCTTGAAGCGGCAATTATATGTGCGCGTTACGCTCAT 754
Db 17775 GGGCGGGGCGCACT--TCTTCTCTGCTCCAGATTTTGAATGTGAATTTCTACCA 17831
QY 755 CAATATCCGCAATCCCATCTCCAAAATCTCACTCTGTGATGACCTTTCCCTACTTAT 814
Db 17832 TGACTTCTTCTCTCATCTCCACTATTTCACTGATGATGACCTTTCTCTACTTAT 17891
QY 815 AGGCTTAGCATGCTGAGGCGCATGACGAGAGGCTGCTCATCTGAGGCGCAT 874
Db 17892 AGGCTTAGCATGCTGAGGCGCATGACGAGAGGCTGCTCATCTGAGGCGCAT 17951
QY 875 CTGATACCATGCTGCGCGCGCGGAGATACCTGTATGATGATGATGATGATGATGATG 934
Db 17952 CTGATACCATGCTGCGCGCGCGGAGATACCTGTATGATGATGATGATGATGATGATG 18009
QY 935 CCTGTCCCTGCTGCGGAGATCTCTGAGTGTGATGATGATGATGATGATGATGATG 994
Db 18010 CCTGTCCCTGCTGCGGAGATCTCTGAGTGTGATGATGATGATGATGATGATGATG 18069
QY 995 TGATTCGTTTGTGAGTGAAGATGATGATGATGATGATGATGATGATGATGATGATG 1054
Db 18070 TGATTCGTTTGTGAGTGAAGATGATGATGATGATGATGATGATGATGATGATGATG 18129
QY 1055 TGTGATCTCTGTGAGTGAAGATGATGATGATGATGATGATGATGATGATGATGATG 1114
Db 18130 TGTGATCTCTGTGAGTGAAGATGATGATGATGATGATGATGATGATGATGATGATG 18189
QY 1115 GATGCGCTGACGAGGCTGTGATGATGATGATGATGATGATGATGATGATGATGATG 1174
Db 18190 GATGCGCTGACGAGGCTGTGATGATGATGATGATGATGATGATGATGATGATGATG 18249
QY 1175 TGGCGCTGCTTGGCATTCATGAGGCGCTGATGATGATGATGATGATGATGATGATG 1234
Db 18250 TGGCGCTGCTTGGCATTCATGAGGCGCTGATGATGATGATGATGATGATGATGATG 18300
QY 1235 CTATTTTGTGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATG 1294
Db 18301 CTATTTTGTGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATG 18355
QY 1295 CCCCATCATTTACTTCTGTGAGGCTCTTTAGGAGCGCTCAAAATAGGAGAACTGAA 1354
Db 18356 CCCCATCATTTACTTCTGTGAGGCTCTTTAGGAGCGCTCAAAATAGGAGAACTGAA 18415
QY 1355 GCTGCTTCCAGAGGCGCTCTGAGAGAGCGCTGAGTGTGATGATGATGATGATGATG 1414
Db 18416 GCTGCTTCCAGAGGCGCTCTGAGAGAGCGCTGAGTGTGATGATGATGATGATGATG 18475
QY 1415 TCTCTAGAGAAACCTGAGAGCTGTGAGAGAGAGATGATGATGATGATGATGATGATG 1474
Db 18476 TCTCTAGAGAAACCTGAGAGCTGTGAGAGAGAGATGATGATGATGATGATGATGATG 18535
QY 1475 CCT---GTGAGAGAGAGCTTTGAGAGAGATGCTGCTGCGCACTTTGACATTTATG 1530
```

```
Db 18536 CTTGTCAGTCAAGCGGCACTTTGAGAGAAACATGCGCTGCGCA--CTTTGACATTTATG 18594
QY 1531 CATTTTCTTGAACCTTCTGCGCTCAGAAATGTCTGAGTGTCTCTCAAGCTTTGAAATG 1590
Db 18595 TGTTTTCTTGAACCTTCTGCGCTCAGAAATGTCTGAGTGTCTCTCAAGCTTTGAAATG 18653
QY 1591 ATGTTTATCTAACCTGACAGTTGCAATTTTCAACCAATGAGAAAGATTAATGTCAGTAC 1650
Db 18654 ATGTTTATCTAACCTGACAGTTGCAATTTTCAACCAATGAGAAAGATTAATGTCAGTAC 18712
QY 1651 ATGTTTATCTAACCTGATTAATACCAATCATTTTCTGTTATCTTGAATGATCT 1710
Db 18713 ATGTTTATCTAACCTGATTAATACCAATCATTTTCTGTTATCTTGAATGATCT 18772
QY 1711 TTCTACTGAAACATTTTTC--TGACATTTGATGATTAATTAAGAGTGTCTGCAAA 1769
Db 18773 TTCTACTGAAACATTTTTC--TGACATTTGATGATTAATTAAGAGTGTCTGCAAA 18832
QY 1770 CCTTAAA--CTCTTCTTATACTGTCTTCTGATGATGATGATGATGATGATGATGATG 1828
Db 18833 CTGAGAAAACCTTTTATTAATCTATATTCACCGGACAGATATC--CAAGGAAATTCCT 18891
QY 1829 TATTAATCTGACAGTATGTTTCCCGGAAATGATGATGATGATGATGATGATGATGATG 1888
Db 18892 TATCTTCTCTGAGCTATGTTTCACTTAAACCAATGTTTCTCTTATGATTAAGGATAT 18951
QY 1889 TACTGACATGAGAGCTCAATTTCTTAATGATGATGATGATGATGATGATGATGATGATG 1948
Db 18952 TACTGACATGAGAGAGCTCAATTTCTTAATGATGATGATGATGATGATGATGATGATG 19011
QY 1949 ATTTCAATTAATTAAGCAAAATTAATGATGATGATGATGATGATGATGATGATGATG 2006
Db 19012 ATTTCAATTAATTAAGCAAAATTAATGATGATGATGATGATGATGATGATGATGATG 19071
QY 2007 AGCTTGAAGAAATGTTTATTAATGATGATGATGATGATGATGATGATGATGATGATG 2066
Db 19072 AGCTTGAAGAAATGTTTATTAATGATGATGATGATGATGATGATGATGATGATGATG 19131
QY 2067 ATGTTGCGCTGCTGTTTCTGATTAAGCAATGATGATGATGATGATGATGATGATGATG 2126
Db 19132 ATGTTGCGCTGCTGTTTCTGATTAAGCAATGATGATGATGATGATGATGATGATGATG 19191
QY 2127 TTTCTTATTTAGCTCTTCTGACTTAATGATGATGATGATGATGATGATGATGATGATG 2186
Db 19192 TTTCTTATTTAGCTCTTCTGACTTAATGATGATGATGATGATGATGATGATGATGATG 19250
QY 2187 AGTTTCAATTTGGGCAATGAGAACTCAATTTGATTTTAATGAGAAATTAATCTGAAA 2246
Db 19251 AGTTTCAATTTGGGCAATGAGAACTCAATTTGATTTTAATGAGAAATTAATCTGAAA 19309
QY 2247 ACCATTTATTTATCTTACAGATTTCTTCAATGTTGAGAGAAATTTCTATCTTCCAG 2306
Db 19310 ACCATTTATTTATCTTACAGATTTCTTCAATTTTACAGAGAAATTTCTATCTTCCAG 19369
QY 2307 TTTTGTATTAATTTGTTGATGATGATGATGATGATGATGATGATGATGATGATGATG 2365
Db 19370 TTTTGTATTAATTTGTTGATGATGATGATGATGATGATGATGATGATGATGATGATG 19429
QY 2366 AGCAAAATGAGAAATGATGATGATGATGATGATGATGATGATGATGATGATGATGATG 2425
Db 19430 AGCAAAATGAGAAATGATGATGATGATGATGATGATGATGATGATGATGATGATGATG 19489
QY 2426 ACTTACATGTTGATGAGATTTCTGATGATGATGATGATGATGATGATGATGATGATGATG 2485
Db 19490 ACTTACATGTTGATGAGATTTCTGATGATGATGATGATGATGATGATGATGATGATGATG 19549
QY 2486 TGTCTAAGAAATTAAGCAAGCGCAACATATTTGATGATGATGATGATGATGATGATG 2545
Db 19550 TGTCTAAGAAATTAAGCAAGCGCAACATATTTGATGATGATGATGATGATGATGATG 19599
QY 2546 CAATTTAGGAATTTCTGATGATGATGATGATGATGATGATGATGATGATGATGATGATG 2598
Db 19600 CAATTTAGGAATTTCTGATGATGATGATGATGATGATGATGATGATGATGATGATGATG 19652
```

RESULT 14
AX647081
LOCUS AX647081 1369 bp DNA linear PAT 04-MAR-2003
DEFINITION Sequence 1273 from Patent EP1270724.
AX647081
ACCESSION AX647081 GI:28800068
VERSION
KEYWORDS
SOURCE Homo sapiens (human)
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.
REFERENCE
1 Suwa, M., Asai, K., Akiyama, Y. and Aburatani, H.
Guanosine triphosphate-binding protein coupled receptors
Patent: EP 1270724-A 1273 02-JAN-2003;
National Institute of Advanced Industrial Science and Technology
(JIP) / Center for Advanced Science and Technology Incubation, Ltd.
(JIP)
FEATURES
source
1.1369
Location/Qualifiers
/organism="Homo sapiens"
/mol_type="unassigned DNA"
/db_xref="taxon:9606"
201..1169
/note="unnamed protein product"
/codon_start=1
/protein_id="CAD69666.1"
/db_xref="GI:28800069"
/translation="MDSTIPVLGTELPINGRETPCYKOTLSFTGLTCLIVSLVALNG
NAVVLMLGGRMRNAVSIYILNVADPFLSGHIIICSLPLINIHPIISLSPM
TPEYFGLSMLSAISTPERCSILMPWYHCRBRPYSWCVLLMLSLRLSEMP
CDLPLFSAGNVWCETSDPTITAMVPLFCVVLGSSLYLVRIIGSRKMPILRYVTI
LTLVLPVLCGLPGPIQWALFSLHLDKVLPCFVHLVSLFLALNSSNPILYFPVG
SFRORORQWIKVILQALQDTPEDVDEGGMLPQETLELSGRLEQ"
ORIGIN
Query Match 45.6%; Score 1192.6; DB 6; Length 1369;
Best Local Similarity 99.7%; Pred. No. 1.2e-282;
Matches 1195; Conservative 0; Mismatches 4; Indels 0; Gaps 0;
QY 462 TCATGGGTCATGACGAGGAGGTTCTGAGCATGATCAACATCCAGCTTGGGTACA 521
DB 171 TCCAGGGTCATGACGAGGAGGTTCTGAGCATGATCAACATCCAGCTTGGGTACA 230
QY 522 GAACTGACCAATCAACGAGAGTGAAGAGCTCTTGTCTACCAAGACAGACCTTGAAGCTTC 581
DB 231 GAATGACCAATCAACGAGAGTGAAGAGCTCTTGTCTACCAAGACAGACCTTGAAGCTTC 290
QY 582 ACCGGGCTGACGTGATCGTTTCCCTTGTGCGCTGACAGAGAAACGGGTTTGTCTGCG 641
DB 291 ACCGGGCTGACGTGATCGTTTCCCTTGTGCGCTGACAGAGAAACGGGTTTGTCTGCG 350
QY 642 CTCCTGGGCTGCGCATGCGAGAGAGCTGTCTCAATCTACATCTCAACCTGTGCGG 701
DB 351 CTCCTGGGCTGCGCATGCGAGAGAGCTGTCTCAATCTACATCTCAACCTGTGCGG 410
QY 702 GCCGACTTCTCTTCTTAGCGGCGACATTATATGTTGCGCGTTACCGCTCATCAATATC 761
DB 411 GCCGACTTCTCTTCTTAGCGGCGACATTATATGTTGCGCGTTACCGCTCATCAATATC 470
QY 762 CGGCATCCCATCTCCAAATCTCTAGTCTGTGATGACCTTTCCCTACTTTATAGGCTTA 821
DB 471 CGGCATCCCATCTCCAAATCTCTAGTCTGTGATGACCTTTCCCTACTTTATAGGCTTA 530
QY 822 AGCATGCTGAGCGGCATCAGACCGAGCGCTGCTGCATCTGTGAGCCCATCTGTGATC 881
DB 531 AGCATGCTGAGCGGCATCAGACCGAGCGCTGCTGCATCTGTGAGCCCATCTGTGATC 590
QY 882 CACTGCGCGCGCCGACAGATCTGTCAATCGGTCAATGTGTCTGTCTGTGCGCCCTGTGC 941
DB 591 CACTGCGCGCGCCGACAGATCTGTCAATCGGTCAATGTGTCTGTCTGTGCGCCCTGTGC 650

QY 942 CTGCTGCGAGTATACCTGAGATGATGTTCTGTGACTTCTGTATTAGTGTGATTC 1001
DB 651 CTGCTGCGAGTATACCTGAGATGATGTTCTGTGACTTCTGTATTAGTGTGATTC 710
QY 1002 GTTTGTGTGAAACGTGAGATTTCAATTACAATCGCGTGTGTTTATGTGTGTT 1061
DB 711 GTTTGTGTGAAACGTGAGATTTCAATTACAATCGCGTGTGTTTATGTGTGTT 770
QY 1062 CTCTGTGGGTCCAGCGCTGTCTGTGTGTGATGATTCCTGTGTGATTCGCGAAATGCGG 1121
DB 771 CTCTGTGGGTCCAGCGCTGTCTGTGTGTGATGATTCCTGTGTGATTCGCGAAATGCGG 830
QY 1122 CTGACGAGCTGTACGAGACATCCCTGACAGTGTGATTCCTCTCTCTGTGAGCTG 1181
DB 831 CTGACGAGCTGTACGAGACATCCCTGACAGTGTGATTCCTCTCTCTGTGAGCTG 890
QY 1182 CCTTTGGCATTTGAGTGGGCTGTGTTTCCAGATCCACCTGTGATTTGAAAGTCTTATTT 1241
DB 891 CCTTTGGCATTTGAGTGGGCTGTGTTTCCAGATCCACCTGTGATTTGAAAGTCTTATTT 950
QY 1242 TGTCAATGTGATCTAGTTTCATTTTCTGTCCGCTTTTAAACAGCATGCGCAACCCATC 1301
DB 951 TGTCAATGTGATCTAGTTTCATTTTCTGTCCGCTTTTAAACAGCATGCGCAACCCATC 1010
QY 1302 ATTTACTTCTTGTGAGGCTCTTTTAAAGCAGCTCAAAATAGGAGAACTGAAAGCTGTT 1361
DB 1011 ATTTACTTCTTGTGAGGCTCTTTTAAAGCAGCTCAAAATAGGAGAACTGAAAGCTGTT 1070
QY 1362 CTCAGAGGCTCTGACAGACAGCGCTGAGTGAATGAATGAGGAGGTTCTTCAG 1421
DB 1071 CTCAGAGGCTCTGACAGACAGCGCTGAGTGAATGAATGAGGAGGTTCTTCAG 1130
QY 1422 GAAACCTGAGCTGTGCGGAGAGCAATTTGAGCAGTGAAGAAACCTTGTGCTGTCA 1481
DB 1131 GAAACCTGAGCTGTGCGGAGAGCAATTTGAGCAGTGAAGAAACCTTGTGCTGTCA 1190
QY 1482 GACAGACCTTTGAGAGCAATGCTGCCCTGACCTTGAACAATTATATGATTTTCTTA 1541
DB 1191 GACAGACCTTTGAGAGCAATGCTGCCCTGACCTTGAACAATTATATGATTTTCTTA 1250
QY 1542 GCCTTGTGCTCGAATAATGTTCTGAGTGTCTCTCAAGTCTTGAATAGATTTATCTA 1601
DB 1251 GCCTTGTGCTCGAATAATGTTCTGAGTGTCTCTCAAGTCTTGAATAGATTTATCTA 1310
QY 1602 ACCTGACAGTTGCAATTTTACCCATGAGAAAGCAATTAGTGAACAGTACATGTTTGA 1660
DB 1311 ACCTGACAGTTGCAATTTTACCCATGAGAAAGCAATTAGTGAACAGTACATGTTTGA 1369
RESULT 15
AX429467
LOCUS AX429467 8622 bp DNA linear PAT 21-JUN-2002
DEFINITION Sequence 3 from Patent WO0234914.
AX429467
ACCESSION AX429467
VERSION AX429467.1 GI:21540764
KEYWORDS
SOURCE Homo sapiens (human)
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.
REFERENCE
1 Zhao, Q., Beasley, E.M. and Wei, M.H.
Isolated human g-protein coupled receptors, nucleic acid molecules
encoding human gpcr proteins, and uses thereof
Patent: WO 0234914-A 3 02-MAY-2002;
PR CORP NY (US)
FEATURES
source
1..8622
Location/Qualifiers
/organism="Homo sapiens"
/mol_type="unassigned DNA"
/db_xref="taxon:9606"
ORIGIN

Query Match 42.2%; Score 1104; DB 6; Length 8622;
 Best Local Similarity 99.6%; Pred. No. 1.2e-260;
 Matches 1107; Conservative 0; Mismatches 5; Indels 0; Gaps 0;

QY 462 TCATGGGTCATGAGCTGGGGTTTCTGACATGATGATTTCAACATCCCATTTGGGTACA 521
 DB 7498 TCCAGGGTCATGACACTGGGGTTTCTGACATGATGATTTCAACATCCCATTTGGGTACA 7557
 QY 522 GAACCTGACCAATCAACGAGCTGAGAGACTCTTGTGTAAGAGACCCCTGAGCTTC 581
 DB 7558 GAACCTGACCAATCAACGAGCTGAGAGACTCTTGTGTAAGAGACCCCTGAGCTTC 7617
 QY 582 ACGGGGCTGACGTGATCTTTTCTTGTGCGCTGACAGAAACGCGGTGTCTG 641
 DB 7618 ACGGGGCTGACGTGATCTTTTCTTGTGCGCTGACAGAAACGCGGTGTCTG 7677
 QY 642 CTCCCTGGGCTGGCGGATGAGGAGAAAGGCTGCTCATATCAATCTCAACCTGGTGG 701
 DB 7678 CTCCCTGGGCTGGCGGATGAGGAGAAAGGCTGCTCATATCAATCTCAACCTGGTGG 7737
 QY 702 GCCGACTTCTCTTCTTATGCGGCAATTAATGTTGCGCGTTACGCTCATCAATATC 761
 DB 7738 GCCGACTTCTCTTCTTATGCGGCAATTAATGTTGCGCGTTACGCTCATCAATATC 7797
 QY 762 CGCCATCCCATCTCAAAAATCTCTAGTCTGTGATGACCTTTCCTTATATAGGCTTA 821
 DB 7798 CGCCATCCCATCTCAAAAATCTCTAGTCTGTGATGACCTTTCCTTATATAGGCTTA 7857
 QY 822 AGCANTGAGAGGCGCATGAGAACGAGGCGCTGCTGCATCTGAGGCGCATCTGGTAC 881
 DB 7858 AGCANTGAGAGGCGCATGAGAACGAGGCGCTGCTGCATCTGAGGCGCATCTGGTAC 7917
 QY 882 CACTGCGCGCGCGCCAGATACCTGTATCATGTGATGCTGCTGCGGCGCTGTCC 941
 DB 7918 CACTGCGCGCGCGCCAGATACCTGTATCATGTGATGCTGCTGCGGCGCTGTCC 7977
 QY 942 CTGCTGCGGAGTATCTGTGAGTGTGATGATCTTCTGTTTATGAGTGTGATCT 1001
 DB 7978 CTGCTGCGGAGTATCTGTGAGTGTGATGATCTTCTGTTTATGAGTGTGATCT 8037
 QY 1002 GTTTGAGTGAACGTCAGATTTCAATTAATCGCGGTGCTTTTATGAGTGTG 1061
 DB 8038 GTTTGAGTGAACGTCAGATTTCAATTAATCGCGGTGCTTTTATGAGTGTG 8097
 QY 1062 CTCTGTGGGCTGAGCTGGTCTGTGCTGATCAAGATTTCTGTGATCCCGAAGATCCG 1121
 DB 8098 CTCTGTGGGCTGAGCTGGTCTGTGCTGATCAAGATTTCTGTGATCCCGAAGATCCG 8157
 QY 1122 CTGACAGGCTGTACGTGACCATCTCTCAAGTGTGCTGCTCTCTGTGCTG 1181
 DB 8158 CTGACAGGCTGTACGTGACCATCTCTCAAGTGTGCTGCTCTCTGTGCTG 8217
 QY 1182 CCCTTTGGCATTCAGTGGGCGCTGTTTCAAGATTCACCTGATTTGAAAGCTATTT 1241
 DB 8218 CCCTTTGGCATTCAGTGGGCGCTGTTTCAAGATTCACCTGATTTGAAAGCTATTT 8277
 QY 1242 TGTCAATGTGATCTAGTTTCAATTTCTGTGCTCTTAAACAGAGTCAACCCCATC 1301
 DB 8278 TGTCAATGTGATCTAGTTTCAATTTCTGTGCTCTTAAACAGAGTCAACCCCATC 8337
 QY 1302 ATTTACTTCTTGTGGGCTCTTTTAAAGCAGCTCAAAATAGGCAAACTGAACTGTT 1361
 DB 8338 ATTTACTTCTTGTGGGCTCTTTTAAAGCAGCTCAAAATAGGCAAACTGAACTGTT 8397
 QY 1362 CTCGAGAGGCTCTGACAGACAGCTGAGGTGATGAAAGTGGAGGTGCTTCTCAG 1421
 DB 8398 CTCGAGAGGCTCTGACAGACAGCTGAGGTGATGAAAGTGGAGGTGCTTCTCAG 8457
 QY 1422 GAAACCTGAGACTGTGGGAGAGAGATTGAGCAGTGAAGAACTCTGCGCTGTCA 1481
 DB 8458 GAAACCTGAGACTGTGGGAGAGAGATTGAGCAGTGAAGAACTCTGCGCTGTCA 8517

QY 1482 GACAGACTTTTGAAGCAATGTGCTGCCCTTACACCTTTACAAATTAATGATTTTCTTA 1541
 DB 8518 GACAGACTTTTGAAGCAATGTGCTGCCCTTACACCTTTACAAATTAATGATTTTCTTA 8577
 QY 1542 GCCTTCTGCTCAGAAATGTCTCAGTGGTCCC 1573
 DB 8578 GCCTTCTGCTCAGAAATGTCTCAGTGGTCCCC 8609

Search completed: October 29, 2004, 13:55:27
 Job time : 10834 secs

GenCore version 5.1.6
Copyright (c) 1993 - 2004 Compugen Ltd.

OM nucleic - nucleic search, using sw model

Run on: October 29, 2004, 09:21:52 ; Search time 1225 Seconds
(without alignments)
11218.763 Million cell updates/sec

Title: US-09-867-570-1
Perfect score: 2618
Sequence: 1 acaaatgcgcgcgaattcgg.....gaaaaaaaaaaaaaaaaa 2618

Scoring table: IDENTITY_NUC
Gapop 10.0 , Gapept 1.0

Searched: 413486 seqs, 2624710521 residues

Total number of hits satisfying chosen parameters: 8269772

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : N Geneseq_23sep04:*

1: geneseqn1980s:*
2: geneseqn1990s:*
3: geneseqn2000s:*
4: geneseqn2001as:*
5: geneseqn2001bs:*
6: geneseqn2002as:*
7: geneseqn2002bs:*
8: geneseqn2003as:*
9: geneseqn2003bs:*
10: geneseqn2003cs:*
11: geneseqn2003ds:*
12: geneseqn2004s:*

Pred. No. is the number of results predicted by chance to have a
score greater than or equal to the score of the result being printed,
and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	2618	100.0	2618	6	ABK52822
2	1529.6	58.4	1997	10	ADD18121
3	1483.8	56.7	2040	6	ADD33744
4	1483.8	56.7	2040	12	ADH08519
5	1483.8	56.7	2040	12	ADH08519
6	1192.6	45.6	1369	10	ADH08519
7	1104	42.2	8622	6	ABK52823
8	1094.2	41.8	1400	6	ABK52823
9	1094.2	41.8	1400	8	ABK52823
10	1094.2	41.8	1400	12	ADH08534
11	1040.6	39.7	1369	10	ADH08534
12	1013.6	38.7	1604	6	ABK52823
13	1013.6	38.7	1604	12	ADH08536
14	1013.6	38.7	1604	12	ADH08536
15	1012.2	38.7	1369	10	ADH08536
16	969	37.0	969	6	ABK52823
17	969	37.0	969	6	ABK52823
18	969	37.0	969	12	ADH08536
19	967.4	37.0	969	12	ADH08536
20	966	36.9	966	12	ADH08536
21	966	36.9	1683	10	ADH08536

22	965.8	36.9	969	12	ADP29104
23	965.6	36.9	1116	4	AAH49507
24	964.4	36.8	966	12	ADH04605
25	964.2	36.8	969	3	AAH70342
26	949.8	36.3	969	2	AAZ10067
27	938.6	35.9	969	2	AAZ10068
28	856.2	32.7	997	12	ADH16635
29	856.2	32.7	997	12	ADH16635
30	845.8	32.3	909	8	ADH16635
31	837.8	32.0	969	8	AAH12208
32	837.8	32.0	969	10	ADH17729
33	836.2	31.9	969	2	AAH10071
34	834.6	31.9	969	2	AAH12205
35	834.6	31.9	969	10	ADH17725
36	833.6	31.8	969	2	AAZ10069
37	833	31.8	969	3	AAH12204
38	833	31.8	969	6	ABK04874
39	833	31.8	969	8	ACC44768
40	833	31.8	969	10	ADH17724
41	833	31.8	969	12	ADH03081
42	829.8	31.7	969	2	AAZ10072
43	829.8	31.7	969	4	AAH20943
44	829.8	31.7	969	5	AAH07944
45	829.8	31.7	969	6	AAH27494

ALIGNMENTS

RESULT 1	ABK52822	ABK52822 standard; cDNA; 2618 BP.
ID	ABK52822	
XX	ABK52822;	
AC	ABK52822;	
XX	27-AUG-2002 (first entry)	
DT	27-AUG-2002 (first entry)	
XX	CDNA encoding human G-protein coupled receptor (GPCR).	
XX	Human; G-protein coupled; receptor; GPCR; human protease;	
KW	human therapeutic protein; query sequence; search; gene; ss;	
KW	sequence database; non-human transgenic animal; gene therapy;	
KW	chromosome 3.	
XX	Homo sapiens.	
OS	Homo sapiens.	
XX	Key	Location/Qualifiers
FH	Key	447..1460
FT	FT	/*tag= a
FT	FT	/product= "Human G-protein coupled receptor (GPCR)"
XX	W0200234914-A1.	
XX	02-MAY-2002.	
XX	10-OCT-2001; 2001MO-US031592.	
XX	25-OCT-2000; 2000US-00695045.	
XX	31-MAY-2001; 2001US-00867570.	
XX	(PEKE) PE CORP NY.	
XX	Wei M, Zhao Q, Woodage T, Di Francesco V, Beasley EM,	
XX	WPI; 2002-463360/49.	
XX	P-PSDB; AAU97598.	
DR	Novel isolated G-protein coupled receptor peptide useful for treating	
PT	disorder characterized by absence of, in appropriate or unwanted	
PT	expression of the receptor protein, and as immunogens to raise	
PT	antibodies.	
XX	Claim 4; Fig 1; 75pp; English.	

XX The present invention relates to a new G-protein coupled receptor (GPCR) peptide. The invention is useful for identifying a modulator of GPCR and for treating a disease or condition mediated by a human protease. The invention is also useful as models for the development of human therapeutics, for identifying therapeutic proteins, as targets for development of human therapeutic agents, and as query sequence to perform a search against sequence databases to, for e.g., identify other family members of related sequences. The vector of the invention is useful for producing a GPCR protein or peptide, for conducting cell-based assays involving the GPCR protein or its fragment, for identifying GPCR protein mutants whose functions are affected, and to produce non-human transgenic animals. The present nucleic acid sequence represents the human G-protein coupled receptor (GPCR) gene located on chromosome 3. This sequence encodes the human G-protein coupled receptor (GPCR) protein of the invention.

Sequence 2618 BP, 655 A, 637 C, 546 G, 780 T, 0 U, 0 Other;

Query Match 100.0%; Score 2618; DB 6; Length 2618;

Best Local Similarity 100.0%; Pred. No. 0;

Matches 2618; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 AACAAATGCGCGAATTCGGCAACGAGATGAATCTAGTTTAAAGGTGTAGACCT 60
 DB 1 AACAAATGCGCGAATTCGGCAACGAGATGAATCTAGTTTAAAGGTGTAGACCT 60
 QY 61 CCTCCCTCTCTTCT 120
 DB 61 CCTCCCTCTCTTCT 120
 QY 121 CACCAAGATGGAAGCTTCTGAGGCTCTCCAGAGAGAGAGAGAGAGAGAGAGAGAG 180
 DB 121 CACCAAGATGGAAGCTTCTGAGGCTCTCCAGAGAGAGAGAGAGAGAGAGAGAGAG 180
 QY 181 CAGTCTGAGAGCTTATGACGAGTTAAACCAATTCCTTCAATTAATTTCCCACTCCAG 240
 DB 181 CAGTCTGAGAGCTTATGACGAGTTAAACCAATTCCTTCAATTAATTTCCCACTCCAG 240
 QY 241 TATTTCTTTTGAAGATTTGAGATGAATTAATACAGACAGAGAGAGAGAGAGAGAG 300
 DB 241 TATTTCTTTTGAAGATTTGAGATGAATTAATACAGACAGAGAGAGAGAGAGAGAG 300
 QY 301 ATCCCAAGGTCTTCT 360
 DB 301 ATCCCAAGGTCTTCT 360
 QY 361 TCCACCAAGAGATTTAAAGATTTAAAGATTTAAAGATTTAAAGATTTAAAGATTT 420
 DB 361 TCCACCAAGAGATTTAAAGATTTAAAGATTTAAAGATTTAAAGATTTAAAGATTT 420
 QY 421 TGGGATGTCAAACAGATTAAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 480
 DB 421 TGGGATGTCAAACAGATTAAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 480
 QY 481 GGTTCCTGAGATGATTCACCATCCAGTCTTGGGTACAGAACTGACCAATCAAG 540
 DB 481 GGTTCCTGAGATGATTCACCATCCAGTCTTGGGTACAGAACTGACCAATCAAG 540
 QY 541 GACGTGAGAGAGCTCTTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCT 600
 DB 541 GACGTGAGAGAGCTCTTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCT 600
 QY 601 TTTCCCTGCGCGTGAACAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 660
 DB 601 TTTCCCTGCGCGTGAACAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 660
 QY 661 GCAGAGAGAGCTGTCTCATCTCATCTCATCTCATCTGTGTGTGTGTGTGTGTGT 720
 DB 661 GCAGAGAGAGCTGTCTCATCTCATCTCATCTCATCTGTGTGTGTGTGTGTGTGT 720
 QY 721 GCGGCACTTATATGTTGCGGTACGCTCATCATATCCGCATCCCATCTCCAAA 780
 DB 721 GCGGCACTTATATGTTGCGGTACGCTCATCATATCCGCATCCCATCTCCAAA 780

DB 721 GCGGCACTTATATGTTGCGGTACGCTCATCATATCCGCATCCCATCTCCAAA 780
 QY 781 TCTCAGTCTGTGATGACCTTTCCCTAATTATAGGCTTAAGAGAGAGAGAGAGAG 840
 DB 781 TCTCAGTCTGTGATGACCTTTCCCTAATTATAGGCTTAAGAGAGAGAGAGAGAG 840
 QY 841 GCAACGAGAGAGCT 900
 DB 841 GCAACGAGAGAGCT 900
 QY 901 ACCGTGATCGGTGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 960
 DB 901 ACCGTGATCGGTGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 960
 QY 961 AGTGAATGTTCTGATCT 1020
 DB 961 AGTGAATGTTCTGATCT 1020
 QY 1021 ATTTGATTAATTCGCGGTGCTGTTTATATGATGATGATGATGATGATGATGATGAT 1080
 DB 1021 ATTTGATTAATTCGCGGTGCTGTTTATATGATGATGATGATGATGATGATGATGAT 1080
 QY 1081 TCTGCTGTGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 1140
 DB 1081 TCTGCTGTGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 1140
 QY 1141 CCATCCTCTCTCAAGTGTGTCTTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCT 1200
 DB 1141 CCATCCTCTCTCAAGTGTGTCTTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCT 1200
 QY 1201 CCTGTTTTCAGATTCACCTGTGATTTGAAAGTCTTATTTGTCATGTGATCTAGTT 1260
 DB 1201 CCTGTTTTCAGATTCACCTGTGATTTGAAAGTCTTATTTGTCATGTGATCTAGTT 1260
 QY 1261 CCATTTCTCTGCGCTCTTAAACAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 1320
 DB 1261 CCATTTCTCTGCGCTCTTAAACAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 1320
 QY 1321 CCTTGAAGAGAGTCAAAATAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 1380
 DB 1321 CCTTGAAGAGAGTCAAAATAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 1380
 QY 1381 ACAAGCTGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 1440
 DB 1381 ACAAGCTGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 1440
 QY 1441 GAAGAGATTTGAG 1500
 DB 1441 GAAGAGATTTGAG 1500
 QY 1501 TGTGCGCTGACACCTTGAACATTAATATGATTTTCTTGAAGCTTCTGCTCAAGAA 1560
 DB 1501 TGTGCGCTGACACCTTGAACATTAATATGATTTTCTTGAAGCTTCTGCTCAAGAA 1560
 QY 1561 TCTCAGTGTCTCTCAAGTCTTGAATATGATTTATCTAAGAGAGAGAGAGAGAG 1620
 DB 1561 TCTCAGTGTCTCTCAAGTCTTGAATATGATTTATCTAAGAGAGAGAGAGAGAG 1620
 QY 1621 CAACCAAGAGAGATTTAGTGTGATGATGATGATGATGATGATGATGATGATGAT 1680
 DB 1621 CAACCAAGAGAGATTTAGTGTGATGATGATGATGATGATGATGATGATGATGAT 1680
 QY 1681 ACATTTTCCCTGTATCTTGAACATTAATCTTGAAGAGAGAGAGAGAGAGAGAGAG 1740
 DB 1681 ACATTTTCCCTGTATCTTGAACATTAATCTTGAAGAGAGAGAGAGAGAGAGAGAG 1740
 QY 1741 ATTTGATTAAG 1800
 DB 1741 ATTTGATTAAG 1800
 QY 1801 CCGATAGATCAAAAG 1860
 DB 1801 CCGATAGATCAAAAG 1860


```

Qy 1861 TCATGTTCCCTTTATGACCTGAGGCAATTACTGACGTTGGAAGCTCAATCTTATAAGT 1920
Db 1861 TCATGTTCCCTTTATGACCTGAGGCAATTACTGACGTTGGAAGCTCAATCTTATAAGT 1920
Qy 1921 GAGTTCGTACCTCTAATTCATGTAATTCAGATATATAAGCAAAATATAGACCTTA 1980
Db 1921 GAGTTCGTACCTCTAATTCATGTAATTCAGATATATAAGCAAAATATAGACCTTA 1980
Qy 1981 GAGGAGATTTCCCTTCATATATAAACAAGCTCTTAGAAATGGTTTATATAGAGCCCTTC 2040
Db 1981 GAGGAGATTTCCCTTCATATATAAACAAGCTCTTAGAAATGGTTTATATAGAGCCCTTC 2040
Qy 2041 CTGCAATTTGTCACAGCATGCTGACATGTCGCTGGCTGGTTCTAGTAAAGCAATGCTG 2100
Db 2041 CTGCAATTTGTCACAGCATGCTGACATGTCGCTGGCTGGTTCTAGTAAAGCAATGCTG 2100
Qy 2101 GCCCCTTCCCTTGAGAACTGTAAGTCTTATTTAGCTCTTCTGCACTAATGAACTAG 2160
Db 2101 GCCCCTTCCCTTGAGAACTGTAAGTCTTATTTAGCTCTTCTGCACTAATGAACTAG 2160
Qy 2161 TGAGAGGCTTAAATATATGTCACCAAGTTTCAATTTGGCCATTTGGAACCTCAATATG 2220
Db 2161 TGAGAGGCTTAAATATATGTCACCAAGTTTCAATTTGGCCATTTGGAACCTCAATATG 2220
Qy 2221 ATTTTAAAGTGAATATATCTTGAAGAAACATTTATTTCACTTACAGATCTTTCAGTT 2280
Db 2221 ATTTTAAAGTGAATATATCTTGAAGAAACATTTATTTCACTTACAGATCTTTCAGTT 2280
Qy 2281 GTAGAGAAATCTTCACTTCCAGGTTTGTATTAATTTGTGATTTGTAATCTTTCAGT 2340
Db 2281 GTAGAGAAATCTTCACTTCCAGGTTTGTATTAATTTGTGATTTGTAATCTTTCAGT 2340
Qy 2341 TAGTTTATGCTGTTTACATGAGAGCAAACTGAAACATCTGACCTTTCATGACAA 2400
Db 2341 TAGTTTATGCTGTTTACATGAGAGCAAACTGAAACATCTGACCTTTCATGACAA 2400
Qy 2401 TCTCAATTAATGTTATGATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAAT 2460
Db 2401 TCTCAATTAATGTTATGATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAAT 2460
Qy 2461 GACATACATGAACCTGGAATATTTGTCTAAGAAATATAGCCCAACCAATATTT 2520
Db 2461 GACATACATGAACCTGGAATATTTGTCTAAGAAATATAGCCCAACCAATATTT 2520
Qy 2521 GTAAATTCATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTT 2580
Db 2521 GTAAATTCATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTT 2580
Qy 2581 AGGAGATCTGCTGCTGGAAGAAAAAAGAAAAAAGAAAAAAGAAAAAAGAAAAAAG 2618
Db 2581 AGGAGATCTGCTGCTGGAAGAAAAAAGAAAAAAGAAAAAAGAAAAAAGAAAAAAG 2618

```

RESULT 2
ADD18121 standard; DNA; 1997 BP.

ADD18121;
15-JAN-2004 (first entry)
Human G-protein coupled receptor (GPCR) DNA sequence Seq ID7.
G protein coupled receptor; GPCR; signal transduction pathway; G protein;
Alzheimer's disease; Parkinson's disease; diabetes; dwarfism;
colour blindness; retinal pigmentosa; asthma; depression; schizophrenia;
sleeplessness; hypertension; anxiety; stress; renal failure;
cardiovascular disorder; neural disorder; oncology disorder;
immune disorder; neuroprotective; gene therapy; gene; ds.
Homo sapiens.
OS

```

FH Key Location/Qualifiers
FT CDS 169..1215
FT /*tag= a
FT /product= "Human G-protein coupled receptor"
XX MO2003016478-A2.
XX 27-FEB-2003.
XX 15-AUG-2002; 2002MO-US026017.
XX 20-AUG-2001; 2001US-0313658P.
XX 12-SEP-2001; 2001US-0318675P.
XX 30-OCT-2001; 2001US-0340703P.
XX 26-NOV-2001; 2001US-0333417P.
XX 06-DEC-2001; 2001US-0338367P.
XX 06-FEB-2002; 2002US-0355596P.
XX (BRIM ) BRISTOL-MYERS SQUIBB CO.
XX Feder JN, Ramanathan CS, Gopal S, Mantler GA;
XX WPI; 2003-278558/27.
XX P-PSDB; ADD18024.
XX New nucleic acid, useful for manufacturing a medicament for preventing,
XX treating or ameliorating a medical condition e.g., neural disorder.
XX Claim 1; SEQ ID NO 7; 251bp; English.
XX
XX This invention relates to novel G protein coupled receptors (GPCRs) and
XX their encoding nucleotide sequences. Many medically significant
XX biological processes are mediated by proteins participating in signal
XX transduction pathways involving G proteins. GPCRs are one of the largest
XX receptor superfamilies known. These receptors are biologically important
XX and malfunction of these receptors results in diseases such as
XX Alzheimer's, Parkinson's, diabetes, dwarfism, colour blindness, retinal
XX pigmentosa and asthma. They are also involved in depression,
XX schizophrenia, sleeplessness, hypertension, anxiety, stress, renal
XX failure and other cardiovascular, neural, oncology and immune disorders.
XX A modulator of the GPCRs of the invention may have neuroprotective
XX activity whilst the sequences of the invention may be useful for gene
XX therapy. The invention may also be useful for manufacturing a medicament
XX for preventing, treating or ameliorating a medical condition. The present
XX sequence is the DNA sequence encoding a human GPCR of the invention.
XX
XX Sequence 1997 BP; 459 A; 501 C; 436 G; 601 T; 0 U; 0 Other;
XX
XX Query Match 58.4%; Score 1529.6; DB 10; Length 1997;
XX Best Local Similarity 91.3%; Pred. No. 0;
XX Matches 1674; Conservative 0; Mismatches 139; Indels 21; Gaps 4;
Qy 457 AATCCTCATGGGTCATCAGACTGG-GGTTCTGAGCATGATTAACCATCCAGTCTTG 515
Db 133 AACATTAAGGGGTCTATCAGACTGGTGGTTCTGAGCATGATTAACCATCCAGTCTTG 192
Qy 516 GGTACAGAACTGACACCAATTAAGAGACGTGAGAGACTCTCTTCTCAAGCAACCTG 575
Db 193 GTTACAGAACTGACACCAATTAAGAGACGTGAGAGACTCTCTTCTCAAGCAACCTG 252
Qy 576 AGCTTACAGGGGCTGACGTGATGTTTCCCTTCTGCGCTGACAGAAAGCGGTTGTG 635
Db 253 AGCTTACAGGGGCTGACGTGATGTTTCCCTTCTGCGCTGACAGAAAGCGGTTGTG 312
Qy 636 CTCTGACTCTGGGCTGCGGATGCGAGAAAGCGTCTTCATCTCAATCTCAACCTG 695
Db 313 CTCTGACTCTGGGCTGCGGATGCGAGAAAGCGTCTTCATCTCAATCTCAACCTG 372
Qy 696 GTCCGCGCCCACTTCTCTTCTTAAGGCGCACTTATATGTTCCCGTTAAGGCTCATC 755
Db 373 GTCCGCGCCCACTTCTCTTCTTCTTCCAGATTAATTAATGCTGCTGTACTCATG 432
Qy 756 AATATCGGCATCCCATCTCCAAA-----TCTCATGCTCTGTGATGACCTTT 803

```

```

Db      433 AACCTCTTCTGTTCAATCCATCAATTTCCCTAGCTTCTTCAACCACTGGAAGACTGCT 492
Qy      804 CCTACTTTATATAGGCTTAAGCATGTGAGGAGCCATCAGACCGAGGCGCTGCTCATC 863
Db      493 GCTTACTTGGACGGCTTGAGCATGCTGAGACCGCTGACACGAGGCGCTGCTGCTGCT 552
Qy      864 CTGTGGCCCATCTGTGATCGATGCGCGCGCCCGCCAGATCTGTCACTGGTCACTGTGTC 923
Db      553 CTGTGGCCCATCTGTGATCGATGCGCGCGCCCGCCAGACCTGTCACTGGTCACTGTGTC 612
Qy      924 CTGTCTGTGGGCGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 983
Db      613 CTGTCTGTGGGCGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 672
Qy      984 TTTAGTGTGCTGATTTCTGTTTGTGTGAAAGTCAATTTCAATTCAGTGGGCTG 1043
Db      673 TTTAGTGTGATGTGATCTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGT 732
Qy      1044 GTTTTTTATATGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGT 1103
Db      733 ATTTTATTTATTCATGGTCTCTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGT 792
Qy      1104 GGATTCGGGAAATATGCGGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGT 1163
Db      793 GGCTCAGAGGGGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGT 852
Qy      1164 TTCTCTCTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGT 1223
Db      853 TTCTCTCTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGT 912
Qy      1224 GATTGAAAAGTCTTATTTTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGT 1283
Db      913 GATTCTGATGTCTTATTTTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGT 972
Qy      1284 AGCAATGCCAACCCCATCTTACTTCTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGT 1337
Db      973 AGCAATGCCAACCCCATCTTACTTCTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGT 1032
Qy      1338 AATAGGCGAAGACCTGAGCTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGT 1397
Db      1033 CAGCAGCGCATCTCAAGCTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGT 1092
Qy      1398 GAAAGTGAAGGGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGT 1457
Db      1093 GAAAGTGAAGGGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGT 1152
Qy      1458 TGAAGAAAGACCTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGT 1517
Db      1153 TGAAGAAAGACCTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGT 1212
Qy      1518 TGACAATTAATGCAATTTTCTTAAAGCTTCTGAGCAAAATGTCTGAGTGTCTCTCA 1577
Db      1213 TGACAATTAATGCAATTTTCTTAAAGCTTCTGAGCAAAATGTCTGAGTGTCTCTCA 1272
Qy      1578 GGTCTTGAATAGATGTTTATCTTAACCTGACAGTTGACAGTTTCAACCATGAGAACAT 1637
Db      1273 GGTCTTGAATAGATGTTTATCTTAACCTGACAGTTGACAGTTTCAACCATGAGAACAT 1332
Qy      1638 AGTCTGACAGTACATGTTTGTGATTTCTGTGATTAATCAATACATTTTCCCTGTATTC 1697
Db      1333 AGTCTGACAGTACATGTTTGTGATTTCTGTGATTAATCAATACATTTTCCCTGTATTC 1392
Qy      1698 TTGCACTGATCTTCTCTCACTGAAACATTTTCTGCACTTTTCAATGTATTAAGAGAGT 1757
Db      1393 TTGCACTGATCTTCTCTCACTGAAACATTTTCTGCACTTTTCAATGTATTAAGAGAGT 1452
Qy      1758 TGCTGTCCACCAACCTTAAATCTTCTTATTAATCTTGTCTCACTGATAGTCAAAA 1817
Db      1453 TGCTGTCCACCAACCTTAAATCTTCTTATTAATCTTGTCTCACTGATAGTCAAAA 1512
Qy      1818 GGAAGATTCCTTATTAATCTGTGACAGTATGTTCCCTGAAATCATGTTCCCTTTATG 1877

```

```

Db      1513 GGAAGATTCCTTATTAATCTGTGACAGTATGTTCCCTGAAATCATGTTCCCTTTATG 1572
Qy      1878 ACTGAGGACATTAATGCAATTTGAGAGCTCAATTTCTTAATAAGAGTTCTGCTACTTA 1937
Db      1573 ACTGAGGACATTAATGCAATTTGAGAGCTCAATTTCTTAATAAGAGTTCTGCTACTTA 1632
Qy      1938 AATTCATTGAATTTCTGAGATTAATAAGCAAAATATAGCTTTAGAGAGATTTCTCCCTT 1997
Db      1633 AATTCATTGAATTTCTGAGATTAATAAGCAAAATATAGCTTTAGAGAGATTTCTCCCTT 1692
Qy      1998 CATTAATAACAGCTTAGAAATTTGTTTATGAAATGACCTCTGCTGATTTGTCCACAG 2057
Db      1693 CATTAATAACAGCTTAGAAATTTGTTTATGAAATGACCTCTGCTGATTTGTCCACAG 1752
Qy      2058 CATGTGACATGTGGCTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGT 2117
Db      1753 CATGTGACATGTGGCTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGT 1812
Qy      2118 ACTGTGAATTTCTTATTAATGCTTCTCTGAGCTAATGAACTTAGAGAGAGCTTAATAA 2177
Db      1813 ACTGTGAATTTCTTATTAATGCTTCTCTGAGCTAATGAACTTAGAGAGAGCTTAATAA 1872
Qy      2178 TGTCCACACAGTTTCAATTTTGGCCATTTGAAACCTCAATATTGATTTTAAAGTGAAT 2237
Db      1873 TGTCCACACAGTTTCAATTTTGGCCATTTGAAACCTCAATATTGATTTTAAAGTGAAT 1932
Qy      2238 ATCTT--GAAACCATTTATTAATTCATTACGA 2269
Db      1933 ATCTTGTTTTTCATTTATTAATTCACGGAAGA 1966

RESULT 3
AAD33744
ID AAD33744 standard; DNA; 2040 BP.
XX
AC AAD33744;
XX
DT 01-UTL-2002 (first entry)
XX
DE Human MrgX1 (mas-related gene) DNA.
XX
KW Human; mas-related gene; G-protein coupled receptor; drg-12 protein;
receptor; sensory perception; pain; analgesic; MrgX1; gene; ds.
XX
OS Homo sapiens.
XX
FH Key Location/Qualifiers
FT CDS 328..1296
FT FT /tag=a
FT FT /product="Human MrgX1 protein"
XX
PN W020018355-A2.
XX
PD 08-NOV-2001.
XX
PE 04-MAY-2001; 2001WO-US014519.
XX
PR 04-MAY-2000; 2000US-0202027P.
PR 01-AUG-2000; 2000US-0222344P.
PR 03-NOV-2000; 2000US-00704707.
PR 19-APR-2001; 2001US-0285493B.
XX
PA (CALY ) CALIFORNIA INST OF TECHNOLOGY.
XX
PI Anderson DJ, Dong X, Zylka M, Han S, Simon M;
XX
DR WPI, 2002-171346/22.
DR P-PSDB; AAE21288.
XX
PT Isolated polypeptide, Mrg, which is a G-protein coupled receptor and an
isolated polypeptide, drg-12, which is also a receptor, useful for
identifying agonists or antagonists for treating pain.
XX

```

PS Disclosure; Page 113-115; 185bp; English.

CC The invention relates to Mrg (mas-related gene) protein, which is a G-protein coupled receptor and drug-12 protein, which is a receptor. The invention is useful for identifying compounds that bind to it, especially agonists or antagonists. Administration of an agent (e.g., the identified agonist) that increases the expression of Mrg in a mammal may be used for treating impaired sensory perception in a mammal, especially pain. The CC antagonist may also be useful for treating impaired sensory perception in a mammal. The present sequence is human MrgX1 DNA

XX Sequence 2040 BP; 455 A; 517 C; 457 G; 611 T; 0 U; 0 Other;

Query Match 56.7%; Score 1483.8; DB 6; Length 2040;

Best Local Similarity 91.4%; Pred. No. 0;

Matches 1596; Conservative 0; Mismatches 147; Indels 4; Gaps 2;

Qy 462 TCATGGGTCATCAGACGCGGGTTCTGAGCATGATTCACCATCCAGTCTGGGTACA 521
Cc |||
Db 298 TCAGGGGTCAACAGCTAGGGTTCTGAGCATGATTCACCATCCAGTCTGGGTACA 357
Qy 522 GAATGACACCAATCAACGAGACGAGAGACTCTTCTCTACAGACAGACCTTGAAGTTC 581
Cc |||
Db 358 GAATGACACCAATCAACGAGACGAGAGACTCTTCTCTACAGACAGACCTTGAAGTTC 417
Qy 582 ACGGGGCTGAGTGATGTTCTTCCCTGTCGGGCTGACAGGAAACGGGTTGTGCTGAG 641
Cc |||
Db 418 ACGGTGCTGAGTGATGTTCTTCCCTGTCGGGCTGACAGGAAACGGTGTGCTGAG 477
Qy 642 CTCCTGGGCTGCGGATCGAGAGAACGCTGTCTCATCTACATCTTCAACCTGTGTCGG 701
Cc |||
Db 478 CTCCTGGGCTGCGGATCGAGAGAACGCTGTCTCATCTACATCTTCAACCTGTGTCGG 537
Qy 702 GCCGATCTTCTCTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCT 761
Cc |||
Db 538 GCAAGCTTCTCTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCT 597
Qy 762 CGGATCCCATCTCTCAAAATCTCTGATCTGATGATGATCTTCTTCTTCTTCTTCTTCT 821
Cc |||
Db 598 CCGATTCATCTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCT 657
Qy 822 AGCATGCTGAGCGGCATCAGACCGAGCGCTGCTCTCATCTCTGTCGATCTGATCTGATC 881
Cc |||
Db 658 AGCTTCTGATGCTGCTGATGAGACCGAGCGCTGCTCTCATCTCTGTCGATCTGATCTGATC 717
Qy 882 CACTGCG 941
Cc |||
Db 718 CGCTGCGCACCG 777
Qy 942 CTGCTGCGGAGTATCTCTGAGATGATCTCTGATCTCTCTGATCTCTCTGATCTCTCTG 1001
Cc |||
Db 778 CTGCTGCGGAGATCTCTGAGATGATCTCTGATCTCTCTGATCTCTCTGATCTCTCTG 837
Qy 1002 GTTGTGTGTGAAGCTCAGATTTCTATCAATTCGCGTGGCTGTTTTTTTATGTGTGTT 1061
Cc |||
Db 838 GCTTGTGTGTGAAGCTCAGATTTCTATCAATTCGCGTGGCTGTTTTTTTATGTGTGTT 897
Qy 1062 CTCTGT 1121
Cc |||
Db 898 CTCTGT 957
Qy 1122 CTGACACAGGCTGTACGATCTCTCTCAAGCTGCTGCTCTCTCTCTCTCTCTCTCTCT 1181
Cc |||
Db 958 CTGACACAGGCTGTACGATCTCTCTCAAGCTGCTGCTCTCTCTCTCTCTCTCTCTCT 1017
Qy 1182 CCGTTTGGCATTCAGTGGGCTGTTTCTCAGATCCACCTGATTTGGAAGTCTTATTT 1241
Cc |||
Db 1018 CCGTTTGGCATTCAGTGGGCTGTTTCTCAGATCCACCTGATTTGGAAGTCTTATTT 1077
Qy 1242 TGTCAATGTCAATAGTTTCAATTTCTGTCTCTCTTCAACAGATGTCACACCCATC 1301
Cc |||
Db 1078 TGTCAATGTCAATAGTTTCAATTTCTGTCTCTCTTCAACAGATGTCACACCCATC 1137

Qy 1302 ATTACTCTCTCTGAGGCTCTCTTTCAGGACGCTCAAAATAGGACGAACCTGAGCTGTT 1361
Cc |||
Db 1138 ATTACTCTCTCTGAGGCTCTCTTTCAGGACGCTCAAAATAGGACGAACCTGAGCTGTT 1197
Qy 1362 CTCAGAGGGCTCTGACGAGACGCGCTGAGGTGATGAGGTGAGGGTGGCTTCTCTAG 1421
Cc |||
Db 1198 CTCAGAGGGCTCTGACGAGACGCGCTGAGGTGATGAGGTGAGGGTGGCTTCTCTAG 1257
Qy 1422 GAAACCCCTGAGGCTGTGGGGAAGAGATTTGAGAGCTGAGGAAACCTCTGCGCTGTCA 1481
Cc |||
Db 1258 GAAATCTGAGGCTGTGGGGAAGAGATTTGAGAGCTGAGGAAACCTCTGCGCTGTCA 1317
Qy 1482 GACAGACTTTCAGAGACATGCTGCGCCGACCTTGACATTTATATGATTTTCTTA 1541
Cc |||
Db 1318 GACAGACTTTCAGAGACATGCTGCGCCGACCTTGACATTTATATGATTTTCTTA 1377
Qy 1542 GCCTTCTGCTCAGAAATGTCTCAGTGTCTCTCAAGTCTTTCGATATGATTTTCTTA 1601
Cc |||
Db 1378 GCCTTCTGCTCAGAAATGTCTCAGTGTCTCTCAAGTCTTTCGATATGATTTTCTTA 1437
Qy 1602 ACCTGACAGTTCAGTTTTCACCATGGAAGACATTAAGTCTGACAGTACATGTTGGAT 1661
Cc |||
Db 1438 ACCTGACAGTTCAGTTTTCACCATGGAAGACATTAAGTCTGACAGTACATGTTAGAT 1497
Qy 1662 TCTCTGATATTAACCAATATTTCTCTGATCTTCTGACATGATCTTCTCTACAGAA 1721
Cc |||
Db 1498 TCTCTGATATTAACCAATATTTCTCTGATCTTCTGATCTTCTCTCTCTCTCTCTCT 1557
Qy 1722 CACTTTTCTGACCTTTTCAATTTGATTAATAAGAGTGTCTGTCAACACCTTAAACTCT 1781
Cc |||
Db 1558 CACTTTTCTGACCTTTTCAATTTGATTAATAAGAGTGTCTGTCAACACCTTAAACAC- 1616
Qy 1782 TCTTTATATCTGTTTCTCTACCTGATATTAATAAGAGATCTCTTATATCTGCA 1841
Cc |||
Db 1617 TCTTTATATCTGTTTCTCTACCTGATATTAATAAGAGATCTCTTATATCTCTCA 1676
Qy 1842 GACTATGTTCCCTGAAATCATGTTCCCTTTATGATCTGAGGACATTAAGTCTGAG 1901
Cc |||
Db 1677 GACTATGTTCCCTGAAATCATGTTCCCTTTATGATCTGAGGACATTAAGTCTGAG 1736
Qy 1902 AGCTCAATCTTAAATGAGTGTCTGCTACCTCTTAAATTCATGATTTCTGATATTA 1961
Cc |||
Db 1737 AGCTCAATCTTAAATGAGTGTCTGCTACCTCTTAAATTCATGATTTCTGATATTA 1796
Qy 1962 AAGCAAAATATGACCTTGAAGAGATTCCTCCCTTATATAAACAAGCTTGAAGAT 2021
Cc |||
Db 1797 AAGCAAAATATGACCTTGAAGAGATTCCTCCCTTATATAAACAAGCTTGAAGAT 1856
Qy 2022 TTTATGATATGACCTCTCTCTGATTTGTCACAGATGATGATGATGATGATGATGAT 2081
Cc |||
Db 1857 TTTATGATATGACCTCTCTCTGATTTGTCACAGATGATGATGATGATGATGATGAT 1916
Qy 2082 TCTAGTAAAGCAATGCTGCGCCCTTCCCTTGAGAACTGTAAGTCTTATTTAGCTCT 2141
Cc |||
Db 1917 TCTAGTAAAGCAATGCTGCGCCCTTCCCTTGAGAACTTATAGTCTTATTTAGCTCT 1976
Qy 2142 TCTGAGTAAAGCAATGCTGCGCCCTTCCCTTGAGAACTTATAGTCTTATTTAGCTCT 2201
Cc |||
Db 1977 TCTGAGTAAAGCAATGCTGCGCCCTTCCCTTGAGAACTTATAGTCTTATTTAGCTCT 2033
Qy 2202 ATTGAAA 2208
Cc |||
Db 2034 ATTGAAA 2040

RESULT 4
ADH08519
ID ADH08519 standard; DNA; 2040 BP.

XX ADH08519;
AC ADH08519;
DT 25-MAR-2004 (first entry)
XX

DE DNA sequence Mrx1.
XX mas-related gene D; Mrpd; Analgesic; Vulnerary; Ophthalmological;
KM sensory perception; glaucoma; Mrg; ds.
XX Mus musculus.
OS WO2004003133-A1.
PN 08-JAN-2004.
PD 13-MAY-2003; 2003WO-US015004.
PF 26-JUN-2002; 2002US-00183116.
PR (CALY) CALIFORNIA INST OF TECHNOLOGY.
PA Anderson DJ, Dong X, Zylka M, Han S, Simon MI;
XX WPI; 2004-083025/08.
XX P-PSDB; ADH08520.
XX
XX New mas-related gene D polypeptides, useful as therapeutics or in
PT identifying agonists or antagonists that alter pain perception in a
PT mammal for treating impaired sensory perception, e.g. chronic intractable
PT pain or neuropathic pain.
PS Disclosure; SEQ ID NO 15; 220pp; English.
XX
XX The present invention relates to an isolated mas-related gene D (Mrpd)
CC polypeptide. The Mrpd polypeptides are useful as therapeutics or for
CC identifying compounds, i.e. agonists or antagonists, that alter pain
CC perception in a mammal. The compounds are useful for treating impaired
CC sensory perception, e.g. chronic intractable pain or neuropathic pain,
CC promoting wound healing, restoring normal sensitivity following injury,
CC or treating ocular conditions, particularly those associated with
CC pressure such as glaucoma. The Mrp genes or proteins may be used as
CC molecular probes for the detection of cells or tissues related to or
CC involved with sensory perception. The present sequence represents a MrpA
CC (Mrg subfamily) encoding sequence.
XX
SQ Sequence 2040 BP; 455 A; 517 C; 457 G; 611 T; 0 U; 0 Other;
Query Match 56.7%; Score 1483.8; DB 12; Length 2040;
Best Local Similarity 91.4%; Pred. No. 0;
Matches 1596; Conservative 0; Mismatches 147; Indels 4; Gaps 2;
QY 462 TCATGGGTCTATGAGCTGGGGTTTCTGAGCATGATTCACCATCCAGCTTGGGTACA 521
DB 298 TCCAGGGTCCACCACTAGGGTTTCTGAGCATGATTCACCATCCAGCTTGGGTACA 357
QY 522 GAACGACGACCAATCAAGGAGGAGTGGAGAGCTCTTGTACAAAGAGACCTGAGCTTC 581
DB 358 GAACTGACCAATCAAGGAGGAGTGGAGAGCTCTTGTACAAAGAGACCTGAGCTTC 417
QY 582 ACGGGGCTGACGTCATCGTTTCCCTTTCGCGCTGACAGAGAAACGGGTGTCTCTG 641
DB 418 ACGGGTCTGACGTCATCGTTTCCCTTTCGCGCTGACAGAGAAACGGGTGTCTCTG 477
QY 642 CTCCTGGGCTGGCGGACATGCGGAGAAACGCTGTCTCATCTACATCTCTCAACCTTGCGG 701
DB 478 CTCCTGGGCTGGCGGACATGCGGAGAAACGCTGTCTCATCTACATCTCTCAACCTTGCGG 537
QY 702 GCCGACTTCTCTCTTCTTGAAGGCGGACATTAATGTTGCGCGTTAGCGCTCATCAATATC 761
DB 538 GCGAGCTTCTCTCTTCTTGAAGGCGGCGCTTAATATGTTGCGCGTTAGCGCTCATCAATATC 557
QY 762 GCGCATCTCATCTCAAAATCTCTCAATCTCTGATGATGACCTTCTCTCACTTATAGGCTTA 821
DB 598 CCCCATTACCATCTCTAAATCTCTCAATCTCTGATGATGATGATGATGATGATGATGATGAT 657
QY 822 AGCATGCTGAGGCGCATGACGACGAGCGTCTGCTGCTCATCTCTGAGCGCATCTGATC 881

DB 658 AGCTTTCTGAGTCCGCTGAGACCCGAGGCTGCTGCTCCGCTCTGAGCCCATCTGATC 717
QY 882 CACTGCGCGCGCGCCGAGATACCTGTATCTGATGATGATGATGATGATGATGATGATGAT 941
DB 718 CGCTGCAACGCGCCGACACACCTGTCAAGGAGTGTGTGTGTGTGTGTGTGTGTGTGTGT 777
QY 942 CTGTGCGGAGATCTCTGAGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 1001
DB 778 CTGTGCGGAGATCTCTGAGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 837
QY 1002 GTTTGGTGAAGAGTCAATTTCAATTCAGGAGTGTGTTTATGATGATGATGATGATGATGAT 1061
DB 838 GCTTGTGCAATCAATCAATCAATCAATCAATCAATCAATCAATCAATCAATCAATCAATCA 897
QY 1062 CTCTGAGTCCAGGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 1121
DB 898 CTCTGAGTCCAGGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 957
QY 1122 CTGACCAAGCTGTAGTGAACATCTCTCAAGTGTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 1181
DB 958 CTGACCAAGCTGTAGTGAACATCTCTCAAGTGTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 1017
QY 1182 CCCTTTGCAATCAAGTGTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 1241
DB 1018 CCCTTTGCAATCAAGTGTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 1077
QY 1242 TGTATGTCATCTAGTTCATTTTCATTTTCATTTTCATTTTCATTTTCATTTTCATTTTCAT 1301
DB 1078 TGTATGTCATCTAGTTCATTTTCATTTTCATTTTCATTTTCATTTTCATTTTCATTTTCAT 1137
QY 1302 ATTTACTTTTGTGTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 1361
DB 1138 ATTTACTTTTGTGTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 1197
QY 1362 CTCAGAGGCTCTGAGGACAGGCTGAGTGAAGTGAAGTGAAGTGAAGTGAAGTGAAGTGAAGT 1421
DB 1198 CTCAGAGGCTCTGAGGACAGGCTGAGTGAAGTGAAGTGAAGTGAAGTGAAGTGAAGTGAAGT 1257
QY 1422 GAAACCTTGAGCTGTGAGGAGAGATTTGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 1481
DB 1258 GAAACCTTGAGCTGTGAGGAGAGATTTGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 1317
QY 1482 GACAGAGCTTTGAGAGAGATGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 1541
DB 1318 GACAGAGCTTTGAGAGAGATGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 1377
QY 1542 GCTTGTGCTGAGAAATGCTCAGTGTGCTCAGGAGCTTCAATGATGATGATGATGATGATGAT 1601
DB 1378 GCTTGTGCTGAGAAATGCTCAGTGTGCTCAGGAGCTTCAATGATGATGATGATGATGATGAT 1437
QY 1602 ACCTGACAGTTCAGTTTCAACCAATGAAAGATTAATCTGACAGTAAATGTTGGAT 1661
DB 1438 ACCTGACAGTTCAGTTTCAACCAATGAAAGATTAATCTGACAGTAAATGTTGGAT 1497
QY 1662 TCTCTGTAATTAACAATATCTTCTGTTATCTTGAATGATCTTCTCTCTCTCTCTCTCTCT 1721
DB 1498 TCTCTGTAATTAACAATATCTTCTGTTATCTTGAATGATCTTCTCTCTCTCTCTCTCTCT 1557
QY 1722 CACTTTTTCGACCTTCAATGTAATAAAGAGTGTGCTGCAACACCTTAAACTCT 1781
DB 1558 CACTTTTTCGACCTTCAATGTAATAAAGAGTGTGCTGCAACACCTTAAACTCT 1616
QY 1782 TCTTTATCTGTTTCTCTAAGTATGATCAAAAGAGATGATCTTATTAATCTGCTCA 1841
DB 1617 TCTTTATCTGTTTCTCTAAGTATGATCAAAAGAGATGATCTTATTAATCTGCTCA 1676
QY 1842 GACTATGTTCCCTGAAATCATGTTCCCTTTTATGATCTGAGAGATTAATCTGAGTTGGA 1901
DB 1677 GACTATGTTCCCTGAAATCATGTTCCCTTTTATGATCTGAGAGATTAATCTGAGTTGGA 1736
QY 1902 AGCTAAATCTTAATAGAGATGCTGCTACTCAATTAATCTGATTAATCTGATTAATTA 1961
DB 1737 AGCTAAATCTTAATAGAGATGCTGCTACTCAATTAATCTGATTAATCTGATTAATTA 1796

Qy	1962	AAGCAAAATATATGACCTTAGAGAGATTCCTCCCTCATAAAAACAGCTTAGAAATTGG	2021
Db	1797	GAGCAAAATATATGCTTCCTTAGACAGACTCTCTCTCATAAAAACCTCTCACTATTGG	1856
Qy	2022	TTTATGATGATGACCCCTCTCTCTGATATTGTGCACAGCATGTGTGACATTTGGCTTGGTT	2081
Db	1857	TTTTATATAAAAAGCTTCCTCCCTGTCATTTGTTCAAGCATGTGTGATATTTGGCTTGGTT	1916
Qy	2082	TCTATGTTAAAGCAATCGTGGGCCCTTCCCTTTGAGAACTGGTAAGTTCTTATTTAGCTCT	2141
Db	1917	TCTATGTTAAAGCAACTGTGGGCCCTTCCCTTTGAGAACTTTTAAGTGTCTTATTTAGCTCT	1976
Qy	2142	TCCCGACATATGAACTAGTAGAGAGCCTATATAATATGTCCACACAGTTTCATTTTGGCC	2201
Db	1977	TCTCGACATATGAACTAGTAGAGAGCCTATATAATATGTG--CCCAAGTTCTATTTTGGCC	2033
Qy	2202	ATTGAAA 2208	
Db	2034	ATTGAAA 2040	
RESULT 5			
ID	ADO36620	standard; cDNA; 2040 BP.	
XX	ADO36620;		
DT	12-AUG-2004	(first entry)	
XX			
De	Human mas-related gene X1 (Mrx1) encoding cDNA SEQ ID NO:1.		
XX			
KM	haematological disease; cardiovascular disease;		
KM	peripheral nervous system disorder; central nervous system disorder;		
KM	inflammation; respiratory disease; urological disorder; cancer;		
KM	Mas-related gene X1; Mrx1; G protein-coupled receptor; antidiabetic;		
KM	cytostatic; haemostatic; cardiovascular; cardiant; vasotropic;		
KM	antiarrhythmic; antiarteriosclerotic; CNS; antiparkinsonian; nootropic;		
KM	neuroprotective; cerebroprotective; antiinflammatory; antiallergic;		
KM	thymolimetic; immunosuppressive; antiaesthetic; respiratory;		
KM	nephrotoxic; gene therapy; human; gene; ss.		
XX			
OS	Homo sapiens.		
XX			
Key	Location/Qualifiers		
FT	328..1296		
FT	CDS		
FT	/*tag= a		
FT	/product= "mas-related gene X1 (Mrx1) "		
XX			
XX	WO2004042402-A2.		
PN			
PD	21-MAY-2004.		
XX			
PF	22-OCT-2003; 2003WO-EP011686.		
XX			
PR	04-NOV-2002; 2002EP-00024571.		
XX			
PA	(FARB) BAYER HEALTHCARE AG.		
XX			
PI	Golz S, Brueggemeier U, Geerts A, Thiele R;		
XX			
PI	WPI; 2004-400733/37.		
DR	P-PSDB; ADO36621.		
XX			
PT	Screening for therapeutic agents, useful for treating e.g., urological		
PT	disorders, comprises contacting a test compound with a Mas-related gene		
PT	X1 polypeptide and detecting binding of the test compound to the		
PT	polypeptide.		
XX			
PS	Disclosure; SEQ ID NO 1; 124bp; English.		
XX			
CC	The present invention describes a method of screening for therapeutic		
CC	agents useful in treating disease such as haematological disease,		

cardiovascular disease, disorders of the peripheral and central nervous system, inflammation, respiratory diseases, urological disorders or cancer in a mammal. The method comprises contacting a test compound with Mas-related gene XI (Mrgxi) polypeptide and detecting binding of the test compound to GPR3 polypeptide, or determining Mrgxi polypeptide activity at a certain test compound concentration (or in the absence of the test compound) and at a different concentration of the test compound or at the presence of a compound known to be a Mrgxi polypeptide regulator. Also described: (1) diagnosing a disease defined above in a mammal; (2) a pharmaceutical composition for the treatment of the disease above comprising a Mrgxi polypeptide, a Mrgxi polynucleotide, or a therapeutic agent which binds to a Mrgxi polypeptide or which regulates the Mrgxi polypeptide activity such as a small molecule, an RNA molecule, an antisense oligonucleotide, a polypeptide, an antibody, or a ribozyme; and (3) preparation of a pharmaceutical composition useful for treating the above-defined diseases. Mrgxi has antineoplastic, cytostatic, haemostatic, cardiovascular, cardiant, vasotropic, antiarrhythmic, antiarteriosclerotic, CNS, antiparkinsonian, nootropic, neuroprotective, cerebroprotective, antiinflammatory, antiallergic, thymomelic, immunosuppressive, antiasmatic, respiratory and nephrotropic activities, and can be used in gene therapy. The regulators of Mrgxi are useful for preparing a pharmaceutical composition for treating disease such as hematological diseases, cardiovascular disease, disorders of the peripheral and central nervous system, inflammation, respiratory disorders, urological and cancer in a mammal. They are also useful for the regulation of Mrgxi activity in a mammal having the disease. The nucleotide sequences encoding Mrgxi are useful as hybridisation probes, in constructing oligomers for PCR, for chromosome and gene mapping, in the recombinant production of Mrgxi, in generating antisense DNA or RNA and in molecular biology techniques that have not yet been developed. Mrgxi are useful for immunising a mammal to produce polyclonal antibodies and for diagnostic purposes. The present sequence encodes human Mrgxi, which is given in the exemplification of the present invention. Mrgxi is a G protein-coupled receptor (GPCR).

SQ Sequence 2040 BP; 455 A; 517 C; 457 G; 611 T; 0 U; 0 Other;

Query Match	Score	DB	Length
56.7%	1483.8	12	2040

Matches 1596; Conservative 0; Mismatches 147; Indels 4; Gaps 2;

462 TCATGGGTCATCAGACTGGGTTCTGAGCATGATTCAACCATCCAGTCTGGTACA 521

Db 298 TCCAGGGTCACCACTAGGTTCTGAGCATGATCCAACCATCTCAACCTTGACACA 357

522 GAACTGACACCAATCAACGAGCTGAGGAGACTCCTTGCTACAAGCAGACCCTGAGCTTC 581

Db 358 GAACTGACCAATCAACGGAAGTGAAGAGACTCTTTGCTACAAAGCAGACCTTGAGCCTC 417

582 ACGGGCTGACGTGATCGTTCCCTTGTGCGCTGACAGGAACGGGTTGTGCTCTGG 641

Db 418 ACGGTGCTGACGTGCATCGTTCCCTTGTGGGCTGACAGGAACGCAGTTGTGCTCTGG 477

642 CTCCTGGGCTGCCGCA TGGCA GGAACGCTGTCTCCATCTACATCCCTCAACCTGGCG 701

Db 478 CTCCTGGGCTGCCGCAAGCGAAGCCCTTCTCCATCTACATCCCTCACTTGGCCGA 537

002 GCGGACTTCCCTCTCCCTTAGCGGCCACATATAAGTCCCGTACGCCCTCAI CAAATAC /61

DB 538 GCAGACTTCCTCTCCACGCGCCGCTATAATAATCCCTGTAAGCTCAACAGTATC 597

[illegible][illegible]

0 1 2 3 4 5 6 7 8 9 10 11 12 13 14 15 16 17 18 19 20 21 22 23 24 25 26 27 28 29 30 31 32 33 34 35 36 37 38 39 40 41 42 43 44 45 46 47 48 49 50 51 52 53 54 55 56 57 58 59 60 61 62 63 64 65 66 67 68 69 70 71 72 73 74 75 76 77 78 79 80 81 82 83 84 85 86 87 88 89 90 91 92 93 94 95 96 97 98 99

2025-01-01 00:00:00

100

Db 231 GAACTGACACCAATCAACGACGTGAGAGACTCTCTTGTCTCAAGACAGAACCTTAGAGCTTC 290
QY 582 ACCGGGGCTGACGTGATCATGTTTCCCTTGTGGGGTGAAGAAACGGGTGTGCTCGG 641
Db 291 ACCGGGGCTGACGTGATCATGTTTCCCTTGTGGGGTGAAGAAACGGGTGTGCTCGG 350
QY 642 CTCCTGGGGCTGACGTGATCATGTTTCCCTTGTGGGGTGAAGAAACGGGTGTGCTCGG 701
Db 351 CTCCTGGGGCTGACGTGATCATGTTTCCCTTGTGGGGTGAAGAAACGGGTGTGCTCGG 410
QY 702 GCGGACTTCTCTTCTTAAAGGCGCAATTATATGTTGCGCGTTACGCTCATCAATATC 761
Db 411 GCGGACTTCTCTTCTTAAAGGCGCAATTATATGTTGCGCGTTACGCTCATCAATATC 470
QY 762 GCGGACTTCTCTTCTTAAAGGCGCAATTATATGTTGCGCGTTACGCTCATCAATATC 821
Db 471 GCGGACTTCTCTTCTTAAAGGCGCAATTATATGTTGCGCGTTACGCTCATCAATATC 530
QY 822 AGCATGCTGACGCGCATGACGACGCGGCTGCTGCTCATCTGTGGCCCATCTGTATC 881
Db 531 AGCATGCTGACGCGCATGACGACGCGGCTGCTGCTCATCTGTGGCCCATCTGTATC 590
QY 882 CACTGCGCGCGCCCAAGATACCTGTCAATGTTGTTCTGCTTGGGCGCTGTCC 941
Db 591 CACTGCGCGCGCCCAAGATACCTGTCAATGTTGTTCTGCTTGGGCGCTGTCC 650
QY 942 CTGCTGGGGAGTATCTGTGATGATGATGTTCTGTGATCTCTGTTAAGTGTGATCT 1001
Db 651 CTGCTGGGGAGTATCTGTGATGATGATGTTCTGTGATCTCTGTTAAGTGTGATCT 710
QY 1002 GTTTGGTGTGAAGCTGAGATTTTCAATATGCGGTGCTGTTTATATGTTGAT 1061
Db 711 GTTTGGTGTGAAGCTGAGATTTTCAATATGCGGTGCTGTTTATATGTTGAT 770
QY 1062 CTCTGTGGGTCAAGCTGTGCTGTGTCAGAAATCTCTGTGATCTCTGTGATCTCT 1121
Db 771 CTCTGTGGGTCAAGCTGTGCTGTGTCAGAAATCTCTGTGATCTCTGTGATCTCT 830
QY 1122 CTGACCGAGGCTGATCTGTGATCTCTGTGATCTCTGTGATCTCTGTGATCTCT 1181
Db 831 CTGACCGAGGCTGATCTGTGATCTCTGTGATCTCTGTGATCTCTGTGATCTCT 890
QY 1182 CCTTTGGGCAATGATGAGGCGCTGTTTCAAGATCTCTGTGATCTCTGTGATCTCT 1241
Db 891 CCTTTGGGCAATGATGAGGCGCTGTTTCAAGATCTCTGTGATCTCTGTGATCTCT 950
QY 1242 TGTGATGATCTGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 1301
Db 951 TGTGATGATCTGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 1010
QY 1302 ATTTACTTCTTCTGTTGGGCTCTTTAAGGCAAGGCTCAAAATAGGCAAGCTGATG 1361
Db 1011 ATTTACTTCTTCTGTTGGGCTCTTTAAGGCAAGGCTCAAAATAGGCAAGCTGATG 1070
QY 1362 CTCGAGAGGGGCTGACGACGCTGAGGCTGAGGCTGAGGCTGAGGCTGAGGCT 1421
Db 1071 CTCGAGAGGGGCTGACGACGCTGAGGCTGAGGCTGAGGCTGAGGCTGAGGCT 1130
QY 1422 GAAACCTGAGAGCTGTGGGAGAGATGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 1481
Db 1131 GAAACCTGAGAGCTGTGGGAGAGATGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 1190
QY 1482 GACGAGAGCTTGAAG 1541
Db 1191 GACGAGAGCTTGAAG 1250
QY 1542 GCGCTTGGCGCTCAAAATGCTGAGTGTCTCAAGGCTTCAATATGATGATGATGAT 1601
Db 1251 GCGCTTGGCGCTCAAAATGCTGAGTGTCTCAAGGCTTCAATATGATGATGATGAT 1310
QY 1602 ACCTGACAGTTGCAATTTTCAACCATGGAAGCAATATGATGACAGTCAATGTTTGA 1660

Db 1311 ACCTGACAGTTGCAATTTTCAACCATGGAAGCAATATGATGACAGTCAATGTTTGA 1369
RESULT 7
ABK52823
ID ABK52823 standard; DNA; 8622 BP.
XX
AC ABK52823;
XX
DT 27-AUG-2002 (first entry)
XX
DE Genomic DNA encoding human G-protein coupled receptor (GPCR).
XX
KW Human; G-protein coupled; receptor; GPCR; human protease;
KW human therapeutic protein; query sequence; search; gene; ds;
KW sequence database; non-human transgenic animal; gene therapy;
XX chromosome 3.
XX
OS Homo sapiens.
XX
FH Key Location/Qualifiers
FT CDS 4300..8496
FT /*tag= a
FT /product= "Human G-protein coupled receptor (GPCR) "
FT exon 4300..4319
FT /*tag= b
FT /number= 1
FT intron 4320..7502
FT /*tag= c
FT /number= 1
FT exon 7503..8496
FT /*tag= d
FT /number= 2
XX
PN W0200234914-A1.
XX
PD 02-MAY-2002.
XX
PE 10-OCT-2001; 2001WO-US031592.
XX
PR 25-OCT-2000; 2000US-00695045.
PR 31-MAY-2001; 2001US-00867570.
XX
PA (PEKS) PE CORP NY.
XX
PI Wei M, Zhao Q, Woodage T, Di Francesco V, Beasley EM;
XX
DR WPI; 2002-463360/49.
XX
P-PSDB; AAU97598.
XX
PT Novel isolated G-protein coupled receptor peptide useful for treating
PT disorder characterized by absence of, in appropriate or unwanted
PT expression of the receptor protein, and as immunogens to raise
XX antibodies.
XX
PS Claim 4; Fig 3; 75pp; English.
XX
CC The present invention relates to a new G-protein coupled receptor (GPCR)
CC peptide. The invention is useful for identifying a modulator of GPCR and
CC for treating a disease or condition mediated by a human protease. The
CC invention is also useful as models for the development of human
CC therapeutics, for identifying therapeutic proteins, as targets for
CC development of human therapeutic agents, and as query sequence to perform
CC a search against sequence databases to, for e.g., identify other family
CC members of related sequences. The vector of the invention is useful for
CC producing a GPCR protein or peptide, for conducting cell-based assays
CC involving the GPCR protein or its fragment, for identifying GPCR protein
CC mutants whose functions are affected, and to produce non-human transgenic
CC animals. The present nucleic acid sequence represents the human G-protein
CC coupled receptor (GPCR) gene located on chromosome 3. This sequence
CC encodes the human G-protein coupled receptor (GPCR) protein of the
XX invention

SQ Sequence 8622 BP; 2684 A; 1921 C; 1679 G; 2338 T; 0 U; 0 Other;

Query Match 42.2%; Score 1104; DB 6; Length 8622;
 Best Local Similarity 99.6%; Pred. No. 5 4e-298;
 Matches 1107; Conservative 0; Mismatches 5; Indels 0; Gaps 0;

QY 462 TCATGGGTCATCAGCTGGGGTTTCTGAGCATGATTCACCACTCCAGCTTTGGGTACA 521
 DB 7498 TCAGGGTCATCAGCTGGGGTTTCTGAGCATGATTCACCACTCCAGCTTTGGGTACA 7557

QY 522 GAATGACACCAATCAAGGAGAGTACTCTTGTCTACAGAGACCTGAGCTTC 581
 DB 7558 GAATGACACCAATCAAGGAGAGTACTCTTGTCTACAGAGACCTGAGCTTC 7617

QY 582 ACCGGGCTGACGTCATGCTTCCCTTTCGCGCTGACAGAGAAACGGGTTGCTCTG 641
 DB 7618 ACCGGGCTGACGTCATGCTTCCCTTTCGCGCTGACAGAGAAACGGGTTGCTCTG 7677

QY 642 CTCTGGGCTGCGGATGCGAGAGAAACGCTGTCTCCATCTACATCTGCTGCG 701
 DB 7678 CTCTGGGCTGCGGATGCGAGAGAAACGCTGTCTCCATCTACATCTGCTGCG 7737

QY 702 GCCCATCTCTCTTCTTCTTCTAGCGGCAATATATGTCGCGCTTACGCTCATATC 761
 DB 7738 GCCCATCTCTCTTCTTCTTCTAGCGGCAATATATGTCGCGCTTACGCTCATATC 7797

QY 762 CGCCATCCCATCTCAGAAATCTCAGTCTGATGATCCTTCCCTACTTTATAGGCTTA 821
 DB 7798 CGCCATCCCATCTCAGAAATCTCAGTCTGATGATCCTTCCCTACTTTATAGGCTTA 7857

QY 822 AGCATGCTAGGCGCATCAGACCCAGCGCTGCTGTCTCATCTGTGCGCCATCTGTAC 881
 DB 7858 AGCATGCTAGGCGCATCAGACCCAGCGCTGCTGTCTCATCTGTGCGCCATCTGTAC 7917

QY 882 CACTGCCGCGCGCCAGATACCTGTCTCATGTCATGTCATGTCCTGTCTGCGCGCTGTC 941
 DB 7918 CACTGCCGCGCGCCAGATACCTGTCTCATGTCATGTCATGTCCTGTCTGCGCGCTGTC 7977

QY 942 CTGCTGCGGAGATCTCTGAGATGATGTCCTGATCTTCTGTTTGTGTGCTGATTC 1001
 DB 7978 CTGCTGCGGAGATCTCTGAGATGATGTCCTGATCTTCTGTTTGTGTGCTGATTC 8037

QY 1002 GTTTGTGTGAAACGTCAAGATTCATTACATTCGCTGTGCTGTTTGTGTGTGT 1061
 DB 8038 GTTTGTGTGAAACGTCAAGATTCATTACATTCGCTGTGCTGTTTGTGTGTGT 8097

QY 1062 CTCTGTGTGTCTCAGCTGTCTCTGTCTGTCTGAGATTCCTGTCGAGAGATGCG 1121
 DB 8098 CTCTGTGTGTCTCAGCTGTCTCTGTCTGTCTGAGATTCCTGTCGAGAGATGCG 8157

QY 1122 CTGACCAAGCTGTACGTCATCTCTCTCAGAGTCTGTCTCTCTCTGTGTGCTG 1181
 DB 8158 CTGACCAAGCTGTACGTCATCTCTCTCAGAGTCTGTCTCTCTCTGTGTGCTG 8217

QY 1182 CCTTTGTGCAATCAGTGTGCGCTCTGTTTTCAGAGATTCACCTGTGATTTGAAAGTCTTATTT 1241
 DB 8218 CCTTTGTGCAATCAGTGTGCGCTCTGTTTTCAGAGATTCACCTGTGATTTGAAAGTCTTATTT 8277

QY 1242 TGTATGTGCAATCTGTTTCTGATTTCTGTCGCTCTTAAACAGAGTGTGCAACCCCATC 1301
 DB 8278 TGTATGTGCAATCTGTTTCTGATTTCTGTCGCTCTTAAACAGAGTGTGCAACCCCATC 8337

QY 1302 ATTTACTTCTTGTGTGCTCTTCTTGTAGGAGCTCAAAATAGGAGAGAACTGAGCTGTGTT 1361
 DB 8338 ATTTACTTCTTGTGTGCTCTTCTTGTAGGAGCTCAAAATAGGAGAGAACTGAGCTGTGTT 8397

QY 1362 CTCGAGAGGCTCTGACAGACACGCTGAGTGTGATGAAAGTGTGAGGCTTCTCTCAG 1421
 DB 8398 CTCGAGAGGCTCTGACAGACACGCTGAGTGTGATGAAAGTGTGAGGCTTCTCTCAG 8457

QY 1422 GAAACCTGTGAGCTGTCTGAGAGAGATTTGAGAGCATGTGAGAGAACTCTGTGCTGTCA 1481
 DB 8458 GAAACCTGTGAGCTGTCTGAGAGAGATTTGAGAGCATGTGAGAGAACTCTGTGCTGTCA 8517

QY 1482 GACAGACTTTGAGAGCATGCTGCGCCAGCCCTTGACATTTATGATTTTCTTA 1541
 DB 8518 GACAGACTTTGAGAGCATGCTGCGCCAGCCCTTGACATTTATGATTTTCTTA 8577

QY 1542 GCTTCTGCTCTAGAAATGTCTCAGTGTCC 1573
 DB 8578 GCTTCTGCTCTAGAAATGTCTCAGGTTCC 8609

RESULT 8
 AAD33751
 ID AAD33751 standard; DNA; 1400 BP.

AC AAD33751;
 XX
 AC
 XX
 DT 01-JUL-2002 (first entry)
 XX
 DE Human MrgX3 (mas-related gene) DNA.
 XX
 KM Human; mas-related gene; G-protein coupled receptor; drg-12 protein;
 XX
 KM receptor; sensory perception; pain; analgesic; MrgX3; gene; ds.
 XX
 OS Homo sapiens.
 XX
 FT Key Location/Qualifiers
 FT CDS 332..1300
 FT /tag= a
 FT /product= "Human MrgX3 protein"

XX
 EN WO200183555-A2.
 XX
 PD 08-NOV-2001.
 XX
 PF 04-MAY-2001; 2001WO-US014519.
 XX
 PR 04-MAY-2000; 2000US-0202027P.
 PR 01-AUG-2000; 2000US-022344P.
 PR 03-NOV-2000; 2000US-00704707.
 PR 19-APR-2001; 2001US-0285493P.
 XX
 PA (CALY) CALIFORNIA INST OF TECHNOLOGY.
 XX
 PI Anderson DJ, Dong X, Zylka M, Han S, Simon M;
 XX
 DR WPI; 2002-171346/22.
 DR P-PSDB; AAE21286.
 XX
 PT Isolated polypeptide, Mrg, which is a G-protein coupled receptor and an
 PT isolated polypeptide, drg-12, which is also a receptor, useful for
 PT identifying agonists or antagonists for treating pain.
 XX
 PS Disclosure; Page 128-129; 185pp; English.
 XX
 CC The invention relates to Mrg (mas-related gene) protein, which is a G-
 CC protein coupled receptor and drg-12 protein, which is a receptor. The
 CC invention is useful for identifying compounds that bind to it, especially
 CC agonists or antagonists. Administration of an agent (e.g. the identified
 CC agonist) that increases the expression of Mrg in a mammal may be used for
 CC treating impaired sensory perception in a mammal, especially pain. The
 CC antagonist may also be useful for creating impaired sensory perception in
 CC a mammal. The present sequence is human MrgX3 DNA
 CC
 SQ Sequence 1400 BP; 276 A; 375 C; 363 G; 386 T; 0 U; 0 Other;

Query Match 41.8%; Score 1094.2; DB 6; Length 1400;
 Best Local Similarity 99.7%; Pred. No. 1e-295;
 Matches 1096; Conservative 0; Mismatches 3; Indels 0; Gaps 0;

QY 462 TCATGGGTCATCAGCTGGGGTTTCTGAGCATGATTCACCACTCCAGCTTTGGGTACA 521
 DB 302 TCAGGGTCATCAGCTGGGGTTTCTGAGCATGATTCACCACTCCAGCTTTGGGTACA 361

QY 522 GAATGACACCAATCAACGAGAGTGTGAGAGTCTCTTGTCTCAAGACAGACCTGTAGCTTC 581
 DB 362 GAATGACACCAATCAACGAGAGTGTGAGAGTCTCTTGTCTCAAGACAGACCTGTAGCTTC 421
 QY 582 ACGGGGCTGAGCTGATGCTTTCCCTTGTGGGCTGACAGAGAAACGGGTTGTGTCTGG 641
 DB 422 ACGGGGCTGAGCTGATGCTTTCCCTTGTGGGCTGACAGAGAAACGGGTTGTGTCTGG 481
 QY 642 CTCCTGGGCTGGGAGTGGGAGAGAGCTGTCTCAATCTCAATCTCAATCTGTGGG 701
 DB 482 CTCCTGGGCTGGGAGTGGGAGAGAGCTGTCTCAATCTCAATCTCAATCTGTGGG 541
 QY 702 GCCGACTTCTCTTCTTAGGGGACATATATATGTCGGCTTACAGCTCATCAATATC 761
 DB 542 GCCGACTTCTCTTCTTAGGGGACATATATATGTCGGCTTACAGCTCATCAATATC 601
 QY 762 CGCCATCCATCTCCAAATCTCAATCTGTGATGACCTTTCCTTATATAGGCTTA 821
 DB 602 CGCCATCCATCTCCAAATCTCAATCTGTGATGACCTTTCCTTATATAGGCTTA 661
 QY 822 AGCATGCTGAGGCTGATGAGACCGAGGCGTGGCTGATCTGTGGCCCATCTGTATC 881
 DB 662 AGCATGCTGAGGCTGATGAGACCGAGGCGTGGCTGATCTGTGGCCCATCTGTATC 721
 QY 882 CACTGCGGCGCCGACAGATCTGTCAATGATGATCTGTCTGTGGGCGCTGTCC 941
 DB 722 CACTGCGGCGCCGACAGATCTGTCAATGATGATCTGTCTGTGGGCGCTGTCC 781
 QY 942 CTGCTGGGAGTATCTGTGAGTGTGATGATCTGTCTGTATATAGTGTGATCT 1001
 DB 782 CTGCTGGGAGTATCTGTGAGTGTGATGATCTGTCTGTATATAGTGTGATCT 841
 QY 1002 GTTGTGTGAAAGTGAATTTATATATATATATATATATATATATATATATATAT 1061
 DB 842 GTTGTGTGAAAGTGAATTTATATATATATATATATATATATATATATATATAT 901
 QY 1062 CTCTGTGGGCTGACAGCTGTCTGTGTGAGATCTGTGTGATCTGTGTGATCTGTGTG 1121
 DB 902 CTCTGTGGGCTGACAGCTGTCTGTGTGAGATCTGTGTGATCTGTGTGATCTGTGTG 961
 QY 1122 CTGACAGGCTGATGATGATCTGTGTGATGATGATGATGATGATGATGATGATG 1181
 DB 962 CTGACAGGCTGATGATGATCTGTGTGATGATGATGATGATGATGATGATGATG 1021
 QY 1182 CCTTGTGGGCTGATGATGATCTGTGTGATGATGATGATGATGATGATGATGATG 1241
 DB 1022 CCTTGTGGGCTGATGATGATCTGTGTGATGATGATGATGATGATGATGATGATG 1081
 QY 1242 TGTGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATG 1301
 DB 1082 TGTGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATG 1141
 QY 1302 ATTTACTTCTTCTGTGGGCTCTTTAGGAGGCTCAAAATAGGAGAACTGTAGCTGT 1361
 DB 1142 ATTTACTTCTTCTGTGGGCTCTTTAGGAGGCTCAAAATAGGAGAACTGTAGCTGT 1201
 QY 1362 CTCGAGAGGCTCTGAGAGACACCTGTAGAGTGTGAGAGTGTGAGAGTGTGAGAG 1421
 DB 1202 CTCGAGAGGCTCTGAGAGACACCTGTAGAGTGTGAGAGTGTGAGAGTGTGAGAG 1261
 QY 1422 GAAACCTGTGAGTGTGAGAGACACCTGTAGAGTGTGAGAGTGTGAGAGTGTGAGAG 1481
 DB 1262 GAAACCTGTGAGTGTGAGAGACACCTGTAGAGTGTGAGAGTGTGAGAGTGTGAGAG 1321
 QY 1482 GACAGAGCTTTGAGAGAGATGCTGCTGTGAGAGTGTGAGAGTGTGAGAGTGTGAGAG 1541
 DB 1322 GACAGAGCTTTGAGAGAGATGCTGCTGTGAGAGTGTGAGAGTGTGAGAGTGTGAGAG 1381
 QY 1542 GCCTTGTGCTGAGAGATG 1560
 DB 1382 GCCTTGTGCTGAGAGATG 1400

RESULT 9
 AB242595
 ID AB242595 standard; DNA; 1400 BP.
 XX
 AC AB242595;
 XX
 DT 04-MAR-2003 (first entry)
 XX
 DE Human G protein-coupled receptor MrgX3 nucleotide SEQ ID NO:673.
 XX
 KW G protein-coupled receptor; GPCR; antigenic peptide; gene therapy;
 KW G protein-coupled receptor modulator; antibody; immune-related disease;
 KW growth-related disease; cell regeneration-related disease; AIDS; cancer;
 KW immunological-related cell proliferation disease; autoimmune disease;
 KW Alzheimer's disease; atherosclerosis; infection; osteoarthritis; allergy;
 KW osteoporosis; cardiomyopathy; inflammation; Crohn's disease; diabetes;
 KW graft versus host disease; Parkinson's disease; multiple sclerosis; pain;
 KW psoriasis; anxiety; depression; schizophrenia; dementia; memory loss;
 KW mental retardation; epilepsy; asthma; tuberculosis; obesity; nausea;
 KW hypertension; hypotension; renal disorder; rheumatoid arthritis; trauma;
 KW ulcer; gene; ds.
 KW
 OS Homo sapiens.
 OS
 PN WO200261087-A2.
 PN
 PD 08-AUG-2002.
 XX
 XX 19-DEC-2001; 2001WO-US050107.
 XX
 XX 19-DEC-2000; 2000US-0257144P.
 PR
 XX
 XX (LIFE-) LIFESPAN BIOSCIENCES INC.
 PA
 XX Burner GC, Roush CL, Brown JP;
 PI
 XX WPI; 2003-046718/04.
 XX P-PSDB; ABP81750.
 DR
 DR
 PT New isolated antigenic peptides e.g., for G protein-coupled receptors
 PT (GPCR), useful for diagnosing and designing drugs for treating conditions
 PT in which GPCRs are involved, e.g. AIDS, Alzheimer's disease, cancer or
 PT autoimmune diseases.
 PT
 XX
 XX
 PS Disclosure; Fig 1; 523pp; English.
 XX
 XX The present invention describes antigenic peptides (1) comprising: (a)
 CC any one of 1601 sequences (see ABP82019 to ABP83619) of 12-24 amino
 CC acids. Also described: (1) an assay for the detection of a particular G
 CC protein-coupled receptor (GPCR) or a candidate polypeptide in a sample;
 CC and (2) an isolated antibody having high specificity and high affinity or
 CC avidity for a particular GPCR. (1) can be used as GPCR modulators and in
 CC gene therapy. The antigenic peptides for GPCRs are useful in detecting an
 CC antibody against a particular GPCR, and in the production of specific
 CC antibodies. The peptides and antibodies are also useful for detecting the
 CC presence or absence of corresponding GPCRs. The antigenic peptides for
 CC GPCRs and antibodies are useful for diagnosing and designing drugs for
 CC treating immune-related diseases, growth-related diseases, cell
 CC regeneration-related disease, immunological-related cell proliferative
 CC diseases, or autoimmune diseases, e.g. AIDS, Alzheimer's disease,
 CC atherosclerosis, bacterial, fungal, protozoan or viral infections,
 CC osteoarthritis, osteoporosis, cancer, cardiomyopathy, chronic and acute
 CC inflammation, allergies, Crohn's disease, diabetes, graft versus host
 CC disease, Parkinson's disease, multiple sclerosis, pain, psoriasis,
 CC anxiety, depression, schizophrenia, dementia, mental retardation, memory
 CC loss, epilepsy, asthma, tuberculosis, obesity, nausea, hypertension,
 CC hypotension, renal disorders, rheumatoid arthritis, trauma, ulcers, or
 CC any other disorder in which GPCRs are involved. The antibodies may be
 CC used in immunoassays and immunodiagnosis. AB242523 to AB242669 encode
 CC GPCR proteins given in ABP81675 to ABP82018, which are used in the
 CC exemplification of the present invention
 CC
 XX Sequence 1400 BP; 276 A; 375 C; 363 G; 386 T; 0 U; 0 Other;


```
Db 422 ACCGGGGTGAAGTGCATGTTTCCCTTGTGGCGTGAAGAAACGGGTTGTCTCGG 481
Qy 642 CTCCTGGGGCTCCGCGATGCGGAGAAAGCTGTCTTCATCTACATCTCAACCTGTGGCG 701
Db 482 CTCCTGGGGCTCCGCGATGCGGAGAAAGCTGTCTTCATCTACATCTCAACCTGTGGCG 541
Qy 702 GCGGACTTCTCTCTTAGGCGGCACATTATATGTGCGCGTACGCGCTCATCAATATC 761
Db 542 GCGGACTTCTCTCTTAGGCGGCACATTATATGTGCGCGTACGCGCTCATCAATATC 601
Qy 762 CGGCATCCCATCTCCAAATCTCATGCTGTGATGACCTTCCCTACTTATATAGGCTTA 821
Db 602 CGGCATCCCATCTCCAAATCTCATGCTGTGATGACCTTCCCTACTTATATAGGCTTA 661
Qy 822 AGCATGCTGAGCGGCATCAGACCGAGCGCTGCTTCATCTGTGCGCCATCTGGTAC 881
Db 662 AGCATGCTGAGCGGCATCAGACCGAGCGCTGCTTCATCTGTGCGCCATCTGGTAC 721
Qy 882 CACTGCGCGCGCGCGAGATACCTGTCACTGCTCATGATGATGCTGCTGCGCGCTGCTGC 941
Db 722 CACTGCGCGCGCGCGAGATACCTGTCACTGCTCATGATGATGCTGCTGCGCGCTGCTGC 781
Qy 942 CTGCTGCGGAGTATCTGAGTGTGATGATGCTGTGATCTCTGTTAGTGTGCTGATTTCT 1001
Db 782 CTGCTGCGGAGTATCTGAGTGTGATGATGCTGTGATCTCTGTTAGTGTGCTGATTTCT 841
Qy 1002 GTTTGTGTGAAGAGTCAGATTTATTAACAATCCGCGTGTGTTTATATGTGTGTT 1061
Db 842 GTTTGTGTGAAGAGTCAGATTTATTAACAATCCGCGTGTGTTTATATGTGTGTT 901
Qy 1062 CTCGTGCGGATCCAGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 1121
Db 902 CTCGTGCGGATCCAGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 961
Qy 1122 CTGACGAGGCTGTATGCTGACCATCTCTCAAGTGTGTGCTTCTCTCTGTGTGCTG 1181
Db 962 CTGACGAGGCTGTATGCTGACCATCTCTCAAGTGTGTGCTTCTCTCTGTGTGCTG 1021
Qy 1182 CCTTTGGCATTCAGTGGGCTGTGTTTTCAGATTCACCTGATGGAAGTCTTATTT 1241
Db 1022 CCTTTGGCATTCAGTGGGCTGTGTTTTCAGATTCACCTGATGGAAGTCTTATTT 1081
Qy 1242 TGTGATGTGATCTAGTTTCCATTTTCTGTGCGCTCTTAACAGAGGCCAACCCATC 1301
Db 1082 TGTGATGTGATCTAGTTTCCATTTTCTGTGCGCTCTTAACAGAGGCCAACCCATC 1141
Qy 1302 ATTACTTCTTCTGAGGCTCCTTTAGGAGCGCTCAAAATAGGAGAACTGAGCTGTT 1361
Db 1142 ATTACTTCTTCTGAGGCTCCTTTAGGAGCGCTCAAAATAGGAGAACTGAGCTGTT 1201
Qy 1362 CTCAGAGGGCTCTGACAGACAGCCCTGAGGTGATGAAAGGTGAGGGTGTCTCTCAG 1421
Db 1202 CTCAGAGGGCTCTGACAGACAGCCCTGAGGTGATGAAAGGTGAGGGTGTCTCTCAG 1261
Qy 1422 GAAACCTTGAAGCTGTGGGAGAGAGATGAGAGAGAGAGAAACCTGTGCGCTGTCA 1481
Db 1262 GAAACCTTGAAGCTGTGGGAGAGAGATGAGAGAGAGAGAAACCTGTGCGCTGTCA 1321
Qy 1482 GACAGACTTGAAGAGCAATGCTGCCCTGACACCTTGAACAATTATATGCAATTTTCTTA 1541
Db 1322 GACAGACTTGAAGAGCAATGCTGCCCTGACACCTTGAACAATTATATGCAATTTTCTTA 1381
Qy 1542 GCCTTGTGCTCAGAAATG 1560
Db 1382 GCCTTGTGCTCAGAAATG 1400
```

```
RESULT 11
ADCB6444
ID ADCB6444 standard; DNA; 1369 BP.
XX
AC ADCB6444;
```

```
XX 01-JAN-2004 (first entry)
DT Human GPCR gene SEQ ID NO:897.
DE
XX de; gene; human; GPCR;
XX guanosine triphosphate-binding protein coupled receptor; gene therapy.
XX Homo sapiens.
XX EP1270724-A2.
XX
XX 02-JAN-2003.
XX
XX 18-JUN-2002; 2002EP-00013517.
XX
XX 18-JUN-2001; 2001JP-00246789.
XX
XX 18-JUN-2001; 2001JP-00246789.
XX
XX (NAAD-) NAT INST ADVANCED IND SCI & TECHNOLOGY.
XX (ADSC-) CENT ADVANCED SCI & TECHNOLOGY INCUBATIO.
XX
XX Suwa M, Asai K, Akiyama Y, Aburatani H;
XX
XX WPI; 2003-315783/31.
XX
XX P-PSDB; ADCB6445.
XX
XX New polynucleotide, useful for preparing a composition for treating a
XX patient in need of increased or suppressed activity or expression of the
XX guanosine triphosphate-binding protein coupled receptor.
XX
XX Claim 1; SEQ ID NO 897; 28bp; English.
XX
XX The invention relates to a novel polynucleotide encoding a guanosine
XX triphosphate-binding protein coupled receptor (GPCR). A polynucleotide of
XX the invention may have a use in gene therapy. The polynucleotide and
XX polypeptide are useful for preparing a composition for treating a patient
XX in need of increased or suppressed activity or expression of the
XX guanosine triphosphate-binding protein coupled receptor. The
XX polynucleotide sequences shown in ADCB5548-ADCB7616 encode GPCR's of the
XX invention.
XX
XX Sequence 1369 BP; 276 A; 364 C; 337 G; 392 T; 0 U; 0 Other;
XX
XX Query Match 39.7%; Score 1040.6; DB 10; Length 1369;
XX Best Local Similarity 91.7%; Pred. No. 1.1e-280;
XX Matches 1100; Conservative 0; Mismatches 99; Indels 0; Gaps 0;
Qy 462 TCATGGGTCATCAGACTGGGGTTCCTGAGCATGATTCACCATCCAGTCTGGTACA 521
Db 171 TCCAGGGTCACACAGACTAGAGGTTTCTGACATGATCCAACTCACTGAGACACA 230
Qy 522 GAATGACACCAATCAACGAGCGTGAAGAGACTCTTGTCTACAGACAGACCTGAGCTTC 581
Db 231 GAATGACACCAATCAACGAGCGTGAAGAGACTCTTGTCTACAGACAGACCTGAGCTTC 290
Qy 582 ACGGGGCTGAGCTGATGTTTCCCTTGTGCGCTGACAGGAAAGCGGTTGTGCTGG 641
Db 291 ACGGGGCTGAGCTGATGTTTCCCTTGTGCGCTGACAGGAAAGCGGTTGTGCTGG 350
Qy 642 CTCCTGGGCTGCGCATGCGGAGAGAGCTGTCTCAATCTCAATCTCAACCTGTGCGG 701
Db 351 CTCCTGGGCTGCGCATGCGGAGAGAGCTGTCTCAATCTCAATCTCAACCTGTGCGG 410
Qy 702 GCGGACTTCTCTCTTAGGCGGCACATTATATGTGCGCGTACGCGCTCATCAATATC 761
Db 411 GCGGACTTCTCTCTTAGGCGGCACATTATATGTGCGCGTACGCGCTCATCAATATC 470
Qy 762 CGGCATCCCATCTCCAAATCTCATGCTGTGATGACCTTCCCTACTTATATAGGCTTA 821
Db 471 CGGCATCCCATCTCCAAATCTCATGCTGTGATGACCTTCCCTACTTATATAGGCTTA 530
Qy 822 AGCATGCTGAGCGGCATCAGACCGAGCGCTGCTTCATCTGTGCGCCATCTGGTAC 881
```

```

Db 531 AGCTTTGAGTCCGCTGAGCAACGAGCGCTGCTCCGCTCTGTCGCCCATCTGGTAC 590
Qy 882 CACTGCCGCCGCCGAGATACCTGTCATGCTCATGTCATGTCCTGCTGAGCCCTGTC 941
Db 591 CGCTGCCACCGCCGACACACCTGTGAGGGGTGTGTGTCTGTCTGAGGCCCTGTCC 650
Qy 942 CTGCTGCCGAGTATCTGAGAGTGTGTCGTGACTTCCTGTTTAGTGTCTGATTC 1001
Db 651 CTGCTGCCGAGATCTGAGAGTGTGTCGTGACTTCCTGTTTAGTGTCTGATTC 710
Qy 1002 GTTGTGTGAAAGTCAGATTCATTAATGCGGTGCTGTTTGTGTTTAGTGTGTT 1061
Db 711 GCTTGTGTGAAAGTCAGATTCATTAATGCGGTGCTGTTTGTGTTTAGTGTGTT 770
Qy 1062 CTCTGTGGGTCAGAGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 1121
Db 771 CTCTGTGGGTCAGAGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 830
Qy 1122 CTGACCAAGGCTGTGATGTCATCTCTCTGCTGCTGCTGCTGCTGCTGCTGCTG 1181
Db 831 CTGACCAAGGCTGTGATGTCATCTCTCTGCTGCTGCTGCTGCTGCTGCTGCTG 890
Qy 1182 CCCTTTGGGATTCAGTGGGCGCTGTTTTCAGGATGCTGCTGCTGCTGCTGCTGCTG 1241
Db 891 CCCTTTGGGATTCAGTGGGCGCTGTTTTCAGGATGCTGCTGCTGCTGCTGCTGCTG 950
Qy 1242 TGTGATGTGATCTGATGTCATTTTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 1301
Db 951 TGTGATGTGATCTGATGTCATTTTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 1010
Qy 1302 ATTTACTTCTTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 1361
Db 1011 ATTTACTTCTTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 1070
Qy 1362 CTCAGAGAGGCTCTGAGAGCAACGCTGAGTGTGATGAGTGTGAGTGTGAGTGTG 1421
Db 1071 CTCAGAGAGGCTCTGAGAGCAACGCTGAGTGTGATGAGTGTGAGTGTGAGTGTG 1130
Qy 1422 GAAACCTGTGAGCTGTGAGAGCAAGATGTGAGAGTGTGAGAGTGTGAGAGTGTG 1481
Db 1131 GAAATCCGTGAGCTGTGAGAGCAAGATGTGAGAGTGTGAGAGTGTGAGAGTGTG 1190
Qy 1482 GACGAGACTTGTGAGAGCAATGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 1541
Db 1191 GACGAGACTTGTGAGAGCAATGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 1250
Qy 1542 GCTTCTGCTGAGAGTGTGAGTGTGAGTGTGAGTGTGAGTGTGAGTGTGAGTGTG 1601
Db 1251 GCTTCTGCTGAGAGTGTGAGTGTGAGTGTGAGTGTGAGTGTGAGTGTGAGTGTG 1310
Qy 1602 ACCTGACAGTGTGAGTGTGAGTGTGAGTGTGAGTGTGAGTGTGAGTGTGAGTGTG 1660
Db 1311 ACCTGACAGTGTGAGTGTGAGTGTGAGTGTGAGTGTGAGTGTGAGTGTGAGTGTG 1369

```

RESULT 12
AAD33752
ID AAD33752 standard; DNA; 1604 BP.

XX AAD33752;

XX 01-JUL-2002 (first entry)

XX Human MrgX4 (mas-related gene) DNA.

XX Human; mas-related gene; G-protein coupled receptor; drg-12 protein;

XX receptor; sensory perception; pain; analgesic; MrgX4; gene; da.

XX Homo sapiens.

XX Key Location/Qualifiers

FT CDS 433..1401 /*tag= a

```

FT /product= "Human MrgX4 protein"
XX MO200183555-A2.
XX 08-NOV-2001.
XX 04-MAY-2001; 2001MO-US014519.
XX 04-MAY-2000; 2000US-0202027P.
XX 01-AUG-2000; 2000US-022344P.
XX 03-NOV-2000; 2000US-00704707.
XX 19-APR-2001; 2001US-0285493P.
XX (CALY ) CALIFORNIA INST OF TECHNOLOGY.
XX Anderson DJ, Dong X, Zylka M, Han S, Simon M;
XX WPI; 2002-171346/22.
XX P-PSDB; AAE21297.
XX Isolated polypeptide, Mrg, which is a G-protein coupled receptor and an
XX isolated polypeptide, drg-12, which is also a receptor, useful for
XX identifying agonists or antagonists for treating pain.
XX Disclosure; Page 130-132; 185pp; English.
XX The invention relates to Mrg (mas-related gene) protein, which is a G-
XX protein coupled receptor and drg-12 protein, which is a receptor. The
XX invention is useful for identifying compounds that bind to it, especially
XX agonists or antagonists. Administration of an agent (e.g. the identified
XX agonist) that increases the expression of Mrg in a mammal may be used for
XX treating impaired sensory perception in a mammal, especially pain. The
XX antagonist may also be useful for treating impaired sensory perception in
XX a mammal. The present sequence is human MrgX4 DNA
XX
SQ Sequence 1604 BP; 337 A; 420 C; 399 G; 448 T; 0 U; 0 Other;
XX
Query Match 38.7%; Score 1013.6; DB 6; Length 1604;
Best Local Similarity 91.0%; Pred. No. 4,5e-273;
Matches 1090; Conservative 0; Mismatches 104; Indels 4; Gaps 1;
Qy 466 GGTGATGAGTGTGGGTTTCTGAGCATGATGATCAACCATCCAGTCTTGGGTACAGAC 525
Db 407 GGGGACCAAGACTAGAGTTTCTGAGCATGATGATCAACCATCCAGTCTTGGGTACAAAC 466
Qy 526 TGACCAATCAACGAGAGTGTGAGAGTCTTGTCTGTAAGAGAGACCTGAGCTTCAAG 585
Db 467 TGACCAATCAACGAGAGTGTGAGAGTCTTGTCTGTAAGAGAGACCTGAGCTTCAAG 526
Qy 586 GGTGAGTGTGATGTTTCTTGTGCGCTGACAGAAAGCGGTTGTCTGAGCTCC 645
Db 527 TGTGAGTGTGATGTTTCTTGTGCGCTGACAGAAAGCGGTTGTCTGAGCTCC 586
Qy 646 TGGGCTGCGGAGTGTGAGAGAGAGAGTGTGATGATGATGATGATGATGATGATG 705
Db 587 TGGGCTGCGGAGTGTGAGAGAGAGAGTGTGATGATGATGATGATGATGATGATG 646
Qy 706 ACTTCTCTTCTTGTGAGGCGCATTAATGTTGCGCTTACGCTCATTAATACGCC 765
Db 647 ACTTCTCTTCTTGTGAGGCGCATTAATGTTGCGCTTACGCTCATTAATACGCC 706
Qy 766 ATCCATCTCAAAAATCTGAGTCTGTGATGACCTTTCCCTAATTATAGGCTTAAGCA 825
Db 707 ATCCATCTCGCAAAAATCTGAGTCTGTGATGACCTTTCCCTAATTATAGGCTTAAGCA 766
Qy 826 TGTGAGGCGCATGACAGACCGGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 885
Db 767 TGTGAGGCGCATGACAGACCGGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 826
Qy 886 GCGGCGGCGGAGATACCTGATGATGATGATGATGATGATGATGATGATGATGATG 945
Db 827 GCGGCGGCGGAGATACCTGATGATGATGATGATGATGATGATGATGATGATGATG 886

```

Qy	946	TGCGAGATCTGAGAGGAGATGTTCTGTGACTTCTCTTATGAGTGTCTATCTGTTT	1005
Db	887	TGTTTAGATCTCGAAGAGAGTTCTGTACTTCTCTTTTAGTGTCCTATTTCTAGTT	946
Qy	1006	GGTGTAAACGTCAAGATTTCAATTAACATGCGGTGCTGTTTTTTATATGTGTGTTCT	1065
Db	947	GGTGTAAACGTCAAGATTTCAATTAACATGCGGTGCTGTTTTTTATATGTGTGTTCT	1006
Qy	1066	GTGGGTCCAGCTGGTCTGTGCGGTCAAGATTTCTGTGATATCCCGAAGATGCGCGTGA	1125
Db	1007	GTGTTTCCAGCTGTCTGTGCTGTGAGATCTCTGTGATATCCCGAAGATGCGCGTGA	1066
Qy	1126	CCAGGCTGTACGTACCAATCTCTCAACAGTGTGTCTCTCTCTGTGAGCTGCTCT	1185
Db	1067	CCAGGCTGTACGTACCAATCTCTCAACAGTGTGTCTCTCTCTCTGTGAGCTGCTCT	1126
Qy	1186	TTGGCAATCAATGGGCGCTGTTTTCCAGGATTCACCTGGATTGGAAATCTTAATTTGTC	1245
Db	1127	TGGCAATCTGAGGGCGCTTAATTTACAGGAATGCACTGAATTTGAAATCTTAATTTGTC	1186
Qy	1246	ATGTGATCTAGTTTCCATTTTCCGTGTCCGCTCTTAACAGACAGTGCACACCCATCATTT	1305
Db	1187	ATGTTTATCTGTGTTGATATGTCCTGTCTCTCTTAACAGTATGTGCCAACCCATCATTT	1246
Qy	1306	ACTTCTTGTGTGGCTCTTTTAGGCAAGCTCAAAATAGGCAAGCTGAAGCTGTGTTCC	1365
Db	1247	ACTTCTTGTGTGGCTCTTTTAGGCAAGCTCAAAATAGGCAAGCTGAAGCTGTGTTCC	1306
Qy	1366	AGAGGGCTCTSCAGGACACAGCTGAGGTGATGAAGGTGAGGGGTCTTCTCAGAGAA	1425
Db	1307	AGAGGGCTCTSCAGGACACAGCTGAGGTGATGAAGGTGAGGGGTCTTCTCAGAGAA	1366
Qy	1426	CCCTGAGACTGTGGGAAGCAGATTGAGCAGTAGAGAAAGAACTGTGCTCT---GTCA	1481
Db	1367	GCTGTGAGCTGTGGGAAGCAGATTGAGGAGCATAGAGGAGAGCTGTGCTCTCAAGTCA	1426
Qy	1482	GACAGGACTTTGAGAGCAATGTGCGCCGTGCACACCTTGACAAATTAATATGATTTTTCTTA	1541
Db	1427	GACGGGACTTTGAGAGCAACGTGTCTGTGCACACCTTGACAAATTAATATGATTTTTCTTA	1486
Qy	1542	GCTTGTGCTCTCAGAAATGTCTCAGTGTCTCCCTCAGAGCTCTTGAAATGATGTTTATCTA	1601
Db	1487	GCGTTGTGCTCTCAGAAATGTCTCAGTGTATCTCAGAGCTCTTGAAATGATGTTTATCTA	1546
Qy	1602	ACCTGACAGTTGCAAGTTTCAACCATGGAAGCAATTAATCTGACAGATCAATGTTTGG	1659
Db	1547	ACCTGACAGTTGCAAGTTTCAACCATGGAAGCAATTAATCTGACAGATCAATGTTTGG	1604

RESULT 13	
AB242602	
ID	AB242602 standard; DNA; 1604 BP.
XX	
AC	AB242602;
XX	
DT	04-MAR-2003 (first entry)
XX	
DE	Human MxR4 G protein-coupled receptor nucleotide SEQ ID NO:688.
XX	
KW	G protein-coupled receptor; GPCR; antigenic peptide; gene therapy;
KW	G protein-coupled receptor modulator; antibody; immune-related disease;
KW	growth-related disease; cell regeneration-related disease; AIDS; cancer;
KW	immunological-related cell proliferative disease; autoimmune disease;
KW	Alzheimer's disease; atherosclerosis; infection; osteoarthritis; allergy
KW	osteoporosis; cardiomyopathy; inflammation; Crohn's disease; diabetes;
KW	graft versus host disease; Parkinson's disease; multiple sclerosis; pain
KW	parosiasis; anxiety; depression; schizophrenia; dementia; memory loss;
KW	mental retardation; epilepsy; asthma; tuberculosis; obesity; nausea;
KW	hypertension; hypotension; renal disorder; rheumatoid arthritis; trauma;
KW	ulcer; gene; de.
XX	
OS	Homo sapiens.
XX	

PN	WO200261087-A2.
XX	
PD	08-AUG-2002.
XX	
PF	19-DEC-2001; 2001WO-US050107.
XX	
PR	19-DEC-2000; 2000US-0257144P.
XX	
PA	(LIFE-) LIFESPAN BIOSCIENCES INC.
XX	
PI	Burner GC, Roush CL, Brown JF;
XX	
DR	WPI; 2003-046718/04.
DR	P-PsDB; ABP81758.
XX	
PT	New isolated antigenic peptides e.g., for G protein-coupled receptors
PT	(GPCR), useful for diagnosing and designing drugs for treating conditions
PT	in which GPCRs are involved, e.g. AIDS, Alzheimer's disease, cancer or
PT	autoimmune diseases.

Disclosure; Fig 1; 523pp; English.

The present invention describes antigenic peptide(s) (I) comprising: (a) any one of 1601 sequences (see ABP82019 to ABP83619) of 12-24 amino acids. Also described: (1) an assay for the detection of a particular GPCR. Also described: (1) an assay for the detection of a particular G protein-coupled receptor (GPCR), or a candidate polypeptide in a sample; and (2) an isolated antibody having high specificity and high affinity or avidity for a particular GPCR. (1) can be used as GPCR modulators and in gene therapy. The antigenic peptides for GPCRs are useful in detecting an antibody against a particular GPCR, and in the production of specific antibodies. The peptides and antibodies are also useful for detecting the presence or absence of corresponding GPCRs. The antigenic peptides for GPCRs and antibodies are useful for diagnosing and designing drugs for treating immune-related diseases, growth-related diseases, cell regeneration-related disease, immunological-related cell proliferative diseases, or autoimmune diseases, e.g., AIDS, Alzheimer's disease, atherosclerosis, bacterial, fungal, protozoan or viral infections, osteoarthritis, osteoporosis, cancer, cardiomyopathy, chronic and acute inflammation, allergies, Crohn's disease, diabetes, graft versus host disease, Parkinson's disease, multiple sclerosis, pain, psoriasis, anxiety, depression, schizophrenia, dementia, mental retardation, memory loss, epilepsy, asthma, tuberculosis, obesity, nausea, hypertension, hypotension, renal disorders, rheumatoid arthritis, trauma, ulcers, or any other disorder in which GPCRs are involved. The antibodies may be used in immunoassays and immunodiagnosis. ABZ42523 to ABZ42869 encode GPCR proteins given in ABP1675 to ABP82018, which are used in the exemplification of the present invention

SQ Sequence 1604 BP; 337 A; 420 C; 399 G; 448 T; 0 U; 0 Other;

Query Match	38.7%	Score 1013.6	DB 8	length 1604
Best Local Similarity	91.0%	Pred. NC. 4.5e-273		
Matches 1090, Conservative	0	Mismatches 104	Indels 4	Gaps 1

[illegible]

D	b	827	GC	CGCCGCCCCACACACCTGTCAGCGGCTGCTGTGCTCTGCTGCGGCGCTGTCCTCC	886
Q	y	946	TC	CGAGATATCTTGAAGTGAATGTTCTGTGACATTTCTGTATTAGTGTGCTGAATTCGTGTT	1005
D	b	887	TG	TTAGATATCTGAGAGTGAAGATTTCTGTGACTTCTGTATTAGTGTGCTGAATTCGATTT	946
Q	y	1006	GG	TGTGAACACGTCAATATTTCAATATCCGCTGTGCTGTTTATATGTGTGTTCTCT	1055
D	b	947	GG	TGTGAACACGTCAATATTTCAATATCCGCTGTGCTGTTTATATGTGTGTTCTCT	1006
Q	y	1066	GT	GGGCTCAGCGCTGTGCTGTGCTGCTGAGATTTCTGTGGAATCCCGAAGATGCGCTGA	1125
D	b	1007	GT	GTGTTTTCAGCTGTGCTGTGCTGCTGAGATTTCTGTGGAATCCCGAAGATGCGCTGA	1066
Q	y	1126	CC	AGGCTGTACGTGACATATCTCTGACAGTGTGCTGTCTCTCTGTGAGCTGTGCT	1189
D	b	1067	CC	AGGCTGTACGTGACATATCTCTGACAGTGTGCTGTCTCTCTGTGAGCTGTGCT	1126
Q	y	1186	TT	GGCATTCAGTGTGCGCTGTGTTTTCAGAGATTCACCTGTGATGGAAGATCTTAATTTGTC	1245
D	b	1127	TC	GGCATTCAGTGTGCGCTGTGTTTTCAGAGATTCACCTGTGATGGAAGATCTTAATTTGTC	1186
Q	y	1246	AT	GTGCAATCTATGTTTCCATTTTCCGTGCTGCTGCTTAACAGAGTGCACCCCATCATTT	1305
D	b	1187	AT	GTGCAATCTATGTTTCCATTTTCCGTGCTGCTGCTTAACAGAGTGCACCCCATCATTT	1246
Q	y	1306	AC	CTTCTTGTGAGGCTCTTTTATGAGCAGCTCAAAATATGACAGAACCTGAGCTGTCTCC	1365
D	b	1247	AC	CTTCTTGTGAGGCTCTTTTATGAGCAGCTCAAAATATGAGAACCTGAGCTGTCTCC	1306
Q	y	1366	AG	AGAGGCTCTGTCAGAGCACGCTGTGATGATGATGAAGTGTGAGGCTGTCTTCAGAGAA	1425
D	b	1307	AG	AGAGGCTCTGTCAGAGCACGCTGTGATGATGATGAAGTGTGAGGCTGTCTTCAGAGAA	1366
Q	y	1426	CC	CTGAGAGCTGTGCGGGAAGCATTTGAGAGCATGAGGAAGAACTGTGCGCT----GTCA	1481
D	b	1367	GC	CTGAGAGCTGTGCGGGAAGCATTTGAGAGCATGAGGAAGAACTGTGCGCTGTGCAATCA	1426
Q	y	1482	GAC	AGACATTTTGAAGCAATGCTGCGCTGACACCTTGAACAATATATGACATTTTCTTA	1541
D	b	1427	GAC	AGACATTTTGAAGCAATGCTGCGCTGACACCTTGAACAATATATGACATTTTCTTA	1486
Q	y	1542	GC	CTTCTGCTCAGAAATGTCTCAAGTGTCTCCTCAAGGTCTTCAATATGATGTTTATCTA	1601
D	b	1487	GC	CTTCTGCTCAGAAATGTCTCAAGTGTCTCCTCAAGGTCTTCAATATGATGTTTATCTA	1546
Q	y	1602	AC	CTGACATGTTTGCAGTTTTCACCCATGGAAGCATTAATCTGACATGATGATTTTGG	1659
D	b	1547	AC	CTGACATGTTTGCAGTTTTCACCCATGGAAGCATTAATCTGACATGATGATTTTGG	1604
RESULT 15					
ADCC8588					
ID					
ADCC8588 standard; DNA; 1369 BP.					
XX					
ADCC8588;					
XX					
DT					
01-JAN-2004 (first entry)					
XX					
DE					
Human GPCR gene SEQ ID NO:1041.					
XX					
KM					
ds; gene; human; GPCR;					
XX					
KW					
guanosine triphosphate-binding protein coupled receptor; gene therapy.					
XX					
OS					
Homo sapiens.					
XX					
PN					
EPI270724-A2.					
XX					
PD					
02-JAN-2003.					
XX					
FP					
18-JUN-2002; 2002EP-00013517.					
XX					

XX	18-JUN-2001; 2001JP-00246789.
PA	(NAD-) NAT INST ADVANCED IND SCI & TECHNOLOGY.
PA	(ADSC-) CENT ADVANCED SCI & TECHNOLOGY INCUBATIO.
XX	
Pt	Suwa M, Asai K, Akiyama Y, Aburatani H;
XX	
DR	WPI; 2003-315783/31.
DR	P-PADB; ADC86589.
PT	New polynucleotide, useful for preparing a composition for treating a patient in need of increased or suppressed activity or expression of the guanosine triphosphate-binding protein coupled receptor.
CC	
PS	Claim 1; SEQ ID NO 1041; 28pp; English.
XX	
CC	The invention relates to a novel polynucleotide encoding a guanosine triphosphate-binding protein coupled receptor (GPCR). A polynucleotide of the invention may have a use in gene therapy. The polynucleotide and polypeptide are useful for preparing a composition for treating a patient in need of increased or suppressed activity or expression of the guanosine triphosphate-binding protein coupled receptor. The polynucleotide sequences shown in AD85548-ADC87616 encode GPCR's of the invention.
CC	
XX	
SQ	Sequence 1369 BP; 279 A; 364 C; 342 G; 384 T; 0 U; 0 Other;
Query Match	38.7%; Score 1012.2; DB 10; Length 1369;
Best Local Similarity	91.0%; Pred. No. 1e-72;
Matches 1088; Conservative	0; Mismatches 103; Indels 4; Gaps 1;
OY	466 GGATCATGAGCTGGGGTTTCTGAGCATTGATTCACCACTCCAGTCTGGGTACAGAAC 525
Dd	175 GGGGCACACGACTGGGGTTTCTGAGCATTGATTCACCACTCCAGTCTGGGTACAAAC 234
OY	526 TGAACCAATCAACGAGCGTAGAGAGACTCTCGTCTCAAGCAGAACCCTGAGCTCACG 585
Dd	235 TGACAACCAATCAACGAGCGTAGAGAGACTCTCGTCTCAACCAACCTTGAGCTCACG 294
OY	586 GAGTGAAGTGATGGTTTTCCCTTGTCCGCGTGAAGAAAAGCGGTTGTCTTGGCTCC 645
Dd	295 TGCTGAGTGTGATTAATTTCTTCCCTGTCCGAGCTGACAGAAAGCGGTTGTCTTGGCTCC 354
OY	646 TGGGTCGCCGATGCGAGAACGCTGTCTTCATCTACATCTCAACTGATGCGGCGG 705
Dd	355 TGGGCTACCGCATGCGAGAACGCTGTCTTCATCTACATCTCAACTGCGCGCAGAG 414
OY	706 ACTTCCCTCTCTTAGAGGGCCAACTTATATGTTGCGCGTTAAGCCTCATCAATATCCGCC 765
Dd	415 ACTTCCCTCTCTTAGAGTTCCAAATATATGCTTTGCAATTAAGCTCATCAATATACGCC 474
OY	766 ATCCCATCTCAAATAACTCTGAGTCTGCTGATGAGACTTTCCTTACTTTATAGGCTTAAGCA 825
Dd	475 ATCTCATTCGCAAAATCTCTGTTCTGTGATGACCTTTCCTTACTTTATAGGCTTAAGCA 534
OY	826 TGCTGAGCGCCATCAGACCGAGGCTGCTGTTCATCTGTGAGCCCATCTGTATCAACT 885
Dd	535 TGCTGAGCGCCATCAGACCGAGGCTGCTGTCTGTGTGAGCCCATCTGTATCAACT 594
OY	886 GCCGCCGCCGACATACCTGTCACTGGTCACTGTGTGCTTGTCTTGGGCCCTGTGCTCCG 945
Dd	595 GCCGCCGCCGCCACACACTGTCAAGCGGTGCTGTGTCTGTCTGTGGGCCCTGTGCTCCG 654
OY	946 TGCGGAGTATCTGAGAGAGATGTTCTGTGACTTCCGTATTAGAGTGTGATTTCTGTTT 1005
Dd	655 TGTTTATGATCTGAGATGAGATTTCTGTGACTTCTGTTTATAGTGTGATTTCTGTTT 714
OY	1006 GTGTGAAAGCTCAGATTTCTATTAACAATCGCGTGTGTTTATTATGTTGTGTCTCT 1065
Dd	715 GTGTGAAAGCTCAGATTTCTATCCAAATCGCGTGTGATTTTATTATGTTGTGTCTCT 774
OY	1066 GTGGATCAGCTGTGCTGTGTGTGAGATTTCTGTGTGATCCCAGAAATGCGGCTGA 1125

```

Db      775 GTGTTTCCAGCCTGCTCTCTGCTGAGATCCTCTGAGATCCCGAAGATGCCGCTGA 834
Qy      1126 CCAGGCTGTAGGTGACCAATCTCTCAACAGTGTGCTTCTCTCTGTGGCTGCCCC 1185
Db      835 CCAGGCTGTAGGTGACCAATCTCTCAACAGTGTGCTTCTCTCTGTGGCTGCCCC 894
Qy      1186 TTGGATTCAGTGGGCCCTGTTTTCAGAGATCCACTGGATTTGAAAGTCTTATTTTTC 1245
Db      895 TCGGCATTCGTGGGGCCCTTAATTTACAGATGACCTGAATTTGGAAAGTCTTATATATGTC 954
Qy      1246 ATGTGATCTAGTTCCTTCCATTTTCTGTCGCTCTTAACAGCAGTGGCCAAACCCATCATTT 1305
Db      955 ATGTTTATCTGATTTGCAATGCTCTCTCTTAACAGTATGCAACCCATCATTT 1014
Qy      1306 ACTTCTTGTTGGGCTCTCTTTAGGCAAGCTCAAAATAGGCAAACTTGAAGCTGTTCTCC 1365
Db      1015 ACTTCTTGTTGGGCTCTCTTTAGGCAAGCTCAAAATAGGCAAACTTGAAGCTGTTCTCC 1074
Qy      1366 AGAGGGCTCTGCAAGCAAGCCTGAGATGATGAAAGTGAAGGCTTCTCTCAGGAAA 1425
Db      1075 AGAGGGCTCTGCAAGCAAGCCTGAGATGATGAAAGTGAAGGCTTCTCTCAGGAAA 1134
Qy      1426 CCTGAGACTGTGCGGAGGAGATGAGAGCAAGTGAAGGAGAACTCTGCCCC---GTCA 1481
Db      1135 GCTTGAAGCTGTGCGGAGGAGATGAGGAGCAAGTGAAGGAGCTCTGCCCCGTCAAGTCA 1194
Qy      1482 GACAGACTTTGAGAGCAATGTGCTGCTGCCCAACCTTGACAAATATATGCAATTTTCTTA 1541
Db      1195 GACGGGACTTTGAGAGCAACACTGTCTGCCCAACCTTGACAAATATATGCAATTTTCTTA 1254
Qy      1542 GCCTTTCGCTCAGAAATGTCTCAGTGTCTCCTCAAGTCTTGAATAGATGTTATCTTA 1601
Db      1255 GCGTTTCGCTCAGAAATGTCTCAGTGTCTCCTCAAGTCTTCAAAATTAATGTTTATCTTA 1314
Qy      1602 ACCTGACAGTGCAGATTTTCAACCCATGAAAGCATTAAGTCAACAGTACATGTT 1656
Db      1315 ACCTGACAGTGCAGATTTTCAACCCATGAAAGCATTAAGTCTGACAGTACATGTT 1369

```

Search completed: October 29, 2004, 10:54:55
 Job time : 1234 secs

GenCore version 5.1.6
Copyright (c) 1993 - 2004 CompuGen Ltd.

OM nucleic - nucleic search, using sw model

Run on: October 29, 2004, 10:12:37 : Search time 7917 Seconds
(without alignments)
12049.920 Million cell updates/sec

Title: US-09-867-570-1
Perfect score: 2618
Sequence: 1 aacacatgcgcgcgaattcg.....gaaaaaaaaaaaaaaaaa 2618

Scoring table: IDENTITY_NUC
Gapop 10.0 , Gapext 1.0

Searched: 32822875 seqs, 18219865908 residues
Total number of hits satisfying chosen parameters: 65645750

Minimum DB seq length: 0
Maximum DB seq length: 200000000

Post-processing: Minimum Match 0%
Maximum Match 100%

Listing first 45 summaries

Database :

EST:*
1: gb_est1:*
2: gb_est2:*
3: gb_est3:*
4: gb_est4:*
5: gb_est5:*
6: gb_est6:*
7: gb_est7:*
8: gb_est8:*
9: gb_est9:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	709.2	27.1	889	6	CA455045
2	663.2	25.3	883	7	CN835944
3	606	23.1	835	7	CN843633
4	604.4	23.1	764	4	BG198766
5	603.2	23.0	827	7	CN831885
6	598.2	22.8	924	7	CN839634
7	586.8	22.4	688	2	BE439409
8	562.6	21.5	662	1	AV731610
9	548.2	20.9	787	4	BG208126
10	544.6	20.8	825	7	CN843258
11	516	19.7	871	7	CN839125
12	504.6	19.3	953	7	CN839765
13	486.4	18.6	871	4	BG210740
14	483.4	18.5	924	6	CD051096
15	468.8	17.9	1009	7	CN837763
16	453.8	17.3	552	8	AQ392205
17	377.6	14.4	895	8	CN840024
18	366	14.0	716	9	AG183584
19	362.4	13.8	811	8	AQ888076
20	358.8	13.7	503	8	AQ888076
21	347	13.3	2075	8	B74348
22	346.2	13.2	2075	8	AK029369
23	344	13.1	725	6	CD468987
24	336.8	12.9	719	6	CD469292

25	335.4	12.8	620	8	AQ696198
26	324.4	12.4	905	6	CD243595
27	318.6	12.2	644	6	CD469930
28	316.6	12.1	644	6	CD470186
29	310.8	11.9	630	6	CD470076
30	308.8	11.8	634	6	CD471204
31	308.8	11.8	634	6	CD471217
32	308.2	11.8	639	6	CD467510
33	307.2	11.7	637	6	CD467799
34	307.2	11.7	954	7	CD842891
35	304.4	11.6	618	4	BI828553
36	300.6	11.5	634	6	CD535184
37	300.4	11.5	515	1	AL705589
38	300.2	11.5	621	6	CD470214
39	298.8	11.4	699	9	CR190735
40	298.2	11.4	618	6	CD470897
41	296.2	11.3	783	6	CD469424
42	296	11.3	446	8	AQ088761
43	293.6	11.2	765	6	CD469412
44	291.2	11.1	770	6	CD469131
45	291	11.1	777	6	CD467625

ALIGNMENTS

RESULT 1
LOCUS CA455045 889 bp mRNA linear EST 12-NOV-2002
DEFINITION AGENCOURT 10735784 MAFCL Homo sapiens CDNA clone IMAGE:6722551 5', mRNA sequence.
ACCESSION CA455045
VERSION CA455045.1 GI:24905369
KEYWORDS EST.
SOURCE Homo sapiens (human)
ORGANISM Homo sapiens
REFERENCE NIH-MGC http://mgc.ncl.nih.gov/
AUTHORS National Institutes of Health, Mammalian Gene Collection (MGC)
TITLE Unpublished (1999)
JOURNAL Contact: Robert Strausberg, Ph.D.
COMMENT Email: cgabs-remail.nih.gov
Tissue Procurement: Kristi A. Bglund, Ira Pastan
CDNA Library Preparation: Invitrogen Corp
DNA Library Arrayed by: The I.M.A.G.E. Consortium (LNL)
DNA Sequencing by: Agencourt Bioscience Corporation
Clone distribution: MGC clone distribution information can be found through the I.M.A.G.E. Consortium/LNL at:
http://image.llnl.gov
Plate: LLM414285 row: h column: 07
High quality sequence stop: 743.
Location/Qualifiers
1. 889
/organism="Homo sapiens"
/mol_type="mRNA"
/cd_xref="taxon:9606"
/clone="IMAGE:6722551"
/cell_line="ZR-75-1, MCF7, SK-BR-3, MDA-MB-231, hTERT-HME1, LNCaP"
/lab_host="EMDH1.0B"
/clone_lib="MAFCL"
/note="Vector: pCMV-SPORT6; Site_1: EcoRV; Site_2: Not I; Subcloned with brain, liver, lung, kidney and muscle. Directionally cloned. Priming method: Oligo-dT. Average insert size: 1800 bp. Library amplification: 26,000 fold. Kristi A. Bglund, James J. Vincent, Robert Strausberg, Bungkok Lee & Ira Pastan: Discovery of new breast cancer genes encoding membrane and secreted proteins. Manuscript submitted."

ORIGIN

TITLE Creation of genome-wide protein expression libraries using random activation of gene expression
 JOURNAL Nat. Biotechnol. 19 (5), 440-445 (2001)
 MEDLINE 21227151
 PUBMED 11329013
 COMMENT Contact: Scott J. Cain
 Athersys, Inc.
 3201 Carnegie Ave, Cleveland, OH 44115, USA
 Tel: 216 431 9900
 Fax: 216 361 9596
 Email: scain@atersys.com

FEATURES
 source
 High quality sequence stop: 559.
 Location/Qualifiers
 1..764
 /organism="Homo sapiens"
 /mol_type="mRNA"
 /db_xref="taxon:9606"
 /cell_line="HT1080"
 /clone_id="Athersys RAGE library"
 /note="See 'Creation of Genome-wide Protein Expression Libraries using Random Activation of Gene Expression', Nature Biotechnology, in press. Note that even though the cell type indicated is HT1080, since a random activation method was used, these sequence tags are not necessarily expressed in HT1080 under normal circumstances."

ORIGIN

Query Match 23.1%; Score 604.4; DB 4; Length 764;
 Best Local Similarity 89.8%; Pred. No. 1.1e-146;
 Matches 67; Conservative 0; Mismatches 76; Indels 2; Gaps 2;

678 ATCTACATCTCAACCTGTCGCGCCGCACTTCTCTCTTACG-GGCCATTATATG 736
 674 ATCTACATCTCAACCTGTCGCGCCGCACTTCTCTCTTACG-GGCCATTATATG 705
 737 TTGCGCGTACGCGCTCATATAT-CCGCCATCCATCTCCAAATCTCACTCTGGA 795
 704 TTCCCTGTACCTTCACTGATATCCGCCATACATCTTAAATCTCTATCTGTGA 645
 796 TGACCTTCTCTCTTATAGAGCCCTAGAGCTGAGGCGCATGACACCGAGCGCTGCC 855
 644 TGATGTTTCTTCTTCTGCAAGCCCTGAGCTTCTGATGCTGAGCAACGAGCGCTGCC 585
 856 TGTCATCTGTGCGCCATCTGTATCACTGCGCGCCGCCAGATACCTGTATCATGCA 915
 584 TGTCGCTGTGCGCCATCTGTATCACTGCGCGCCGCCAGATACCTGTATCATGCA 525
 916 TGTGTCTGTCTGTGCGCCCTGTCTGCTGCGGAGATATCTGAGTGTATCTGTG 975
 524 TGTGTCTGTCTGTGCGCCCTGTCTGCTGCGGAGATATCTGAGTGTATCTGTG 465
 976 ACTTCTGTAGTGTGCTGATCTGTTGTGTGAAACGACATTCATTATCAATACG 1035
 464 GCTTCTGTAGTGTGCTGATCTGTTGTGTGAAACGACATTCATTATCAATACG 405
 1036 CGT 1095
 404 CGT 345
 1096 TTCTCTGT 1155
 344 TTCTCTGT 285
 1156 TGCTGT 1215
 284 TACTGT 225
 1216 TCCACGT 1275
 224 TCCACGT 165
 1276 CTCTTAACGACAGT 1335

Db 164 CTCTTAACGACAGT 105
 Qy 1336 AAAATAGGACAGAACCTGATGTTCTCCAGAGGCTGTGACGACGCGTGAAGTGG 1395
 Db 104 AAAATAGGACAGAACCTGATGTTCTCCAGAGGCTGTGACGACGCGTGAAGTGG 45
 Qy 1396 ATGAAGTGTGAGGCTGT 1425
 Db 44 ATGAAGTGTGAGGCTGT 15

RESULT 5
 CN831885/c 827 bp mRNA linear EST 02-JUN-2004
 LOCUS AGENCOURT 15669850 NIH_MGC 145 Homo sapiens cDNA clone
 DEFINITION IMAGE:30706617 5', mRNA sequence.
 ACCESSION CN831885
 VERSION CN831885.1 GI:47935638
 KEYWORDS EST.
 SOURCE Homo sapiens (human)
 ORGANISM Homo sapiens
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.

REFERENCE 1 (bases 1 to 827)
 NIH-MGC http://mgc.nci.nih.gov/
 Unpublished (1999)
 CONTACT: Daniela S. Gerhard, Ph.D.
 Office of Cancer Genomics
 National Cancer Institute / NIH
 Bldg. 31 Rm10A07 Bethesda, MD 20892
 Email: cgarbs-remail.nih.gov
 Tissue Procurement: GPCR Consortium
 cDNA Library Preparation: GPCR Consortium
 cDNA Library Arrayed by: The I.M.A.G.E. Consortium (LNL)
 DNA Sequencing by: Agencourt Bioscience Corporation
 Clone distribution: MGC clone distribution information can be
 found through the I.M.A.G.E. Consortium/LNL at:
 http://image.llnl.gov
 Plate: IRB14 row: e column: 10
 High quality sequence stop: 653.
 Location/Qualifiers
 1..827

FEATURES

source
 /organism="Homo sapiens"
 /mol_type="mRNA"
 /db_xref="taxon:9606"
 /clone="IMAGE:30706617"
 /issue_type="mixed"
 /lab_host="DH10B"
 /clone_id="NIH_MGC_145"
 /note="Vector: pCDNA3.1; Site 1: varies by clone; Site 2: varies by clone; ORFs were PCR-amplified and cloned into pCDNA3.1 by the GPCR Consortium. Cloning sites vary by clone and include the following: 5'-EcoRV-XmnI/XhoI-3', 5'-EcoRV-XmnI/NotI-3', EcoRV (TA cloned, non-directional). For information about which gene each clone represents, please visit our anonymous ftp site at
 ftp://image.llnl.gov/image/rearrayed_plates/IRB1.presv.dat
 A Note: this is a NIH_MGC Library."

ORIGIN

Query Match 23.0%; Score 603.2; DB 7; Length 827;
 Best Local Similarity 87.8%; Pred. No. 2.3e-146;
 Matches 69; Conservative 0; Mismatches 88; Indels 9; Gaps 3;

Qy 729 ATTATATGTTGGCGCTTACGCTCATCAATAT-CCGCCATCCATCTCCAAATCTT 784
 Db 800 AATAGCTTCCCATTTACGCTCATCAATATTCAGCCCATCTCATTCGCAAAATCTCT 741
 Qy 785 CAGTCCGT 843
 Db 740 CGTTTCTGT 681

QY	1435	IGTGGGGAAGCAGATTGGAGCAGTGAAGAAACCTCTGCCCCT-----GTCAAGAGAGACT	1490
DB	108	TGTGGGGAAGCAGATTGGGCGCATATAGGAGAGCCTCTGCTTCACTCAGACGAGACT	49
QY	1491	TTGAGAGCAATGCTGCCCTGC	1511
DB	48	TTGAGAGCAACTGTCTCTGC	28
RESULT 7			
LOCUS	BE439409	688 bp	mRNA linear
DEFINITION	HTM1-025F1	Homo sapiens cDNA, mRNA sequence.	EST 25-JUL-2000
ACCESSION	BE439409		
VERSION	BE439409.1	GI:9438891	
KEYWORDS	EST.		
SOURCE	Homo sapiens (human)		
ORGANISM	Homo sapiens		
REFERENCE	Eukarya; Metazoa; Chordata; Craniata; Vertebrate; Euteleostomi; Mammalia; Eutheria; Primates; Catarrhini; Homiidae; Homo.		
AUTHORS	1 (bases 1 to 688)		
TITLE	Gonzalez,P., Epstein,D.L. and Borraes,T.		
JOURNAL	Characterization of gene expression in human trabecular meshwork using single-pass sequencing of 1060 clones		
COMMENT	Invest. Ophthalmol. Vis. Sci. (2000) In press		
FEATURES	Contact: Pedro Gonzalez		
SOURCE	Department of Ophthalmology		
	Duke University		
	Duke Eye Center, Erwin Rd, Box 3802, Durham, NC 27710-3802, USA		
	Tel: 919 681 4085		
	Fax: 919 684 8983		
	Email: pedro.gonzalez@duke.edu.		
ORIGIN	Location/Qualifiers		
	1..688		
	/organism="Homo sapiens"		
	/mol_type="mRNA"		
	/db_xref="taxon:9606"		
	/issue_type="Trabecular meshwork"		
	/clone_id="HTM1"		
Query Match	22.4%;	Score 586.8;	DB 2;
Best Local Similarity	98.1%;	Pred. No. 4.3e-142;	Length 688;
Matches 657;	Conservative 0;	Mismatches 7;	Indels 6;
		Gaps 6;	
QY	457	AATCCATAGGGTATCAGACCTGGGGGTTTGGAGCATGGATTCAACATCCAGCTTGG	516
DB	25	AACGATTAAGGGGTATCAGACTGGGGGTTTGGAGCATGGATTCAACATCCAGCTTGG	84
QY	517	GTACAGAATGACCAATCAACGAGCGTGAAGAGACTCTTGTGTACAGAGACCTTGA	576
DB	85	GTACAGAATGACCAATCAACGAGCGTGAAGAGACTCTTGTGTACAGAGACCTTGA	144
QY	577	GCTTACAGGGGCTGACGTGCATCGTTTCCCTTGTGCGCTGACAGAGAAACGGTTGTGC	636
DB	145	GCTTACAGGGGCTGACGTGCATCGTTTCCCTTGTGCGCTGACAGAGAAACGGTTGTGC	204
QY	637	TCTGGCTCTGGGCGTCGCGCATGCGCAGAGAAACGCTGTCATCTCAATCTTCAACCTGG	696
DB	205	TCTGGCTCTGGGCGTCGCGCATGCGCAGAGAAACGCTGTCATCTCAATCTTCAACCTGG	264
QY	697	TGCGGCGCAGACTTCTTCTTCTTGAAGCGCACATATATATGTTCGCGTTACGCTCTACA	756
DB	265	TGCGGCGCAGACTTCTTCTTCTTGAAGCGCACATATATATGTTCGCGTTACGCTCTACA	324
QY	757	ATATCCGCAATCCATCTTCGAAATCTCTGAGTCTCTGTGATGACCTTTTCCCTACTTTATAG	816
DB	325	ATATCCGCAATCCATCTTCGAAATCTCTGAGTCTCTGTGATGACCTTTTCCCTACTTTATAG	384
QY	817	GCCTAAGCATGTGAGCGGCATACAGACCGAGCGGTGCTGTTCATCTCTGTGGCCCATCT	876
DB	385	GCCTAAGCATGTGAGCGGCATACAGACCGAGCGGTGCTGTTCATCTCTGTGGCCCATCT	444

Oy	877	GGATACACATGCGCGCCGCCCAAGATACATCTGATCGGTCATGTGTCTCTCGGAGCC	936
Dd	445	GGTACACATGCGCGCGCCGCCCAAGATACATCTGATCGGTCATGTGTCTCTCGGAGCC	504
Oy	937	TGTCCCTGCTGCGGAGTATCTCGAGTGGATGTTCTGTGACTTCTCTGTTAAGTGCTG	996
Dd	505	TGTCCTGCTGCGGAGTATCTCGAGTGGATGTTCTGTGACTTCTCTGTTAAGTGCTGA	564
Oy	997	ATTCTGTTGGTGTAAGCGTACAGATTCAATTAACACGGAGGCGGTTTTTTATATG	1056
Dd	565	ATTCTGTTGGTGTAAGCGTACAGATTCAATTAACACGGAGGCGGTTTTTTATATG	622
Oy	1057	TGATTCTGTGTGGGTCTCAGCCTGTGCTCTGCTGTGTCAGAGATTCTGTGTGATCCGGA	1116
Dd	623	TGG-TCTGTGTGGGTCTCAG-CGTGGTCTGCTGTGTCAGGA-TCTGTGTGAT-CCGGA	678
Oy	1117	TGCCGCTGAC	1126
Dd	679	TGCAGCTGAC	688
RESULT_8	AV731610	662 bp	662 bp
LOCUS	AV731610	HTF Homo sapiens cDNA clone HTFANKD08 5', mRNA sequence.	
DEFINITION	AV731610	HTF Homo sapiens cDNA clone HTFANKD08 5', mRNA sequence.	
ACCESSION	AV731610	GI:10849155	
VERSION	AV731610.1	GI:10849155	
KEYWORDS	EST.		
SOURCE	Homo sapiens (human)		
ORGANISM	Homo sapiens		
REFERENCE	Eukaryote; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo. 1 (bases 1 to 662)		
AUTHORS	Gu, Y., Peng, Y., Song, H., Huang, Q., Yang, Y., Gao, G., Xiao, H., Xu, X., Li, N., Qian, B., Liu, F., Qu, J., Gao, X., Cheng, Z., Xu, Z., Zeng, L., Xu, S., Gu, W., Yu, Y., Jia, Y., Fu, G., Ren, S., Zhong, M., Lu, G., Hu, R., Chen, J., Chen, Z., and Han, Z.		
TITLE	Homo sapiens cDNA HTF clones		
JOURNAL	Unpublished (2000)		
COMMENT	Contact: Zeguang Han Chinese National Human Genome Center at Shanghai 351 Guo Shoujing Road, Zhangjiang Hi-Tech Park, Pudong, Shanghai 201203, P. R. China Tel: 86-21-50801919(ex.45) Fax: 86-21-50801922 Email: hanzg@hgc.sh.cn This clone is available at CHGC in Shanghai.		
FEATURES	Location/Qualifiers		
source	1..662 /organism="Homo sapiens" /mol_type="mRNA" /db_xref="taxon:9606" /clone="HTFANKD08" /tissue_type="Hypothalamus" /dev_stage="Adult" /lab_host="SOLR" /clone_1lb="HTF" /note="Vector: pBluescript sk(-); Site_1: EcoRI; Site_2: XhoI"		
ORIGIN			
Query Match	21.5%	Score 562.6;	DB 1; Length 662;
Best Local Similarity	95.5%	Pred. No. 9.3e-136;	
Matches 662;	Conservative	0; Mismatches 25;	Indels 4; Gaps 4;
Oy	1560	GTTCTAGTGTGTCCTCTCAAGGCTCTTTCGAATGATGTTTATCTTAACCTGAGAGTTGCAGTTT	1619
Dd	1	GTTCTAGTGTGTAACCAAGGCTCTCAATGAATGAATGTTTATCTTAACCTGAGAGTTGCAGTTT	60
Oy	1620	TACCCATGGAAGGATTAAGCTAGACAGTCAATAGTTGGATTCCTCGATATTAACCA	1675
Dd	61	TACCCATGGAAGGATTAAGCTAGACAGTCAATAGTTGGATTCCTCGATATTAACCA	120

QY 1680 TACATTTCCCTGTTATCTGACCTGAATCTTTCTACTGAACACTTTTCTGCACCTTT 1739
 Db 121 TACATTTCCCTGTTATCTGACCTGAATCTTTCTACTGAACACTTTTCTGCACCTTT 180
 QY 1740 CATTTGTAATTAAGAGAGTCTGTCACAAACCTTAATACTCTTCTTATATCTGTTTCT 1799
 Db 181 CATTTGTAATTAAGAGAGTCTGTCACAAACCTTAATACTCTTCTTATATCTGTTTCT 240
 QY 1800 ACCGTATAGTATCAAAAAGAGATCTCTTATTAATCTGTCAGACTATGTTCCCTGAAA 1859
 Db 241 ACCGTATAGTATCAAAAAGAGATCTCTTATTAATCTGTCAGACTATGTTCCCTGAAA 300
 QY 1860 ATCATGTTCCCTTTATGACGTGAGGAGTATGACAGTATGAGGAGCTCAATTTCTTAATAG 1919
 Db 301 ATCATGTTCCCTTTATGACGTGAGGAGTATGACAGTATGAGGAGCTCAATTTCTTAATAG 360
 QY 1920 TGAATCTGCTACTCTTAATTCATTAATTCATTAATTCATTAATTCATTAATTCATTAAT 1979
 Db 361 TGAATCTGCTACTCTTAATTCATTAATTCATTAATTCATTAATTCATTAATTCATTAAT 420
 QY 1980 AGAGAGAGATTTCTCCCTTCATTAATAAGATCTTAATAATGTTTATGAAATGACCTCT 2039
 Db 421 AGAGAGAGATTTCTCCCTTCATTAATAAGATCTTAATAATGTTTATGAAATGACCTCT 480
 QY 2040 CCGTATCTGTCACAGCAGTGTGAGTGTGCTGTTGTTTCTAGTAAAGCAATGCT 2099
 Db 481 CCGTATCTGTCACAGCAGTGTGAGTGTGCTGTTGTTTCTAGTAAAGCAATGCT 538
 QY 2100 GGGCCCTTCCCTTGAAGACGTGTAAGTCTTATTAATGCTCTTCTGAGCAATGAATGA 2159
 Db 539 GGGCCCTTCCCTTGAAGACGTGTAAGTCTTATTAATGCTCTTCTGAGCAATGAATGA 597
 QY 2160 GTGAGAG-CCATTAATTAATGTCACCAAGTTCATTTGGCCATTGGA 2208
 Db 598 GTGAGAG-CCATTAATTAATGTCACCAAGTTCATTTGGCCATTGGA 647

RESULT 9
 BG208126/c 787 bp mRNA linear EST 21-APR-2001
 LOCUS RST27616 Atherys RAGE Library Homo sapiens cDNA, mRNA sequence.
 DEFINITION BG208126
 ACCESSION BG208126
 VERSION BG208126.1 GI:13729813
 KEYWORDS EST.
 SOURCE Homo sapiens (human)
 ORGANISM Homo sapiens
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.
 REFERENCE 1 (bases 1 to 787)
 AUTHORS Harrington,J.J., Sherf,B., Rundlett,S., Jackson,P.D., Perry,R., Cain,S., Leventhal,C., Thornton,M., Ramachandran,R., Whittington,J., Lerner,L., Costanzo,D., McElisgot,K., Bozser,S., Mays,R., Smith,B., Veloso,N., Kliska,A., Hees,J., Cothren,K., Lo,K., Offenbacher,J., Danzig,J. and Ducar,M.
 TITLE Creation of genome-wide protein expression libraries using random activation of gene expression
 JOURNAL Nat. Biotechnol. 19 (5), 440-445 (2001)
 MEDLINE 21227151
 PUBMED 11329013
 COMMENT Contact: Scott J. Cain
 Atherys, Inc.
 3201 Carnegie Ave, Cleveland, OH 44115, USA
 Tel: 216 431 9900
 Fax: 216 361 9596
 Email: scain@atherys.com
 High quality sequence stop: 553.
 Location/Qualifiers
 1..787
 /organism="Homo sapiens"
 /mol_type="mRNA"
 /db_xref="taxon:9606"
 /cell_line="HT1080"

/clone lib="Atherys RAGE library"
 /note="See 'Creation of Genome-wide Protein Expression Libraries using Random Activation of Gene Expression', Nature Biotechnology, in press. Note that even though the cell type indicated is HT1080, since a random activation method was used, these sequence tags are not necessarily expressed in HT1080 under normal circumstances."

Query Match 20.9%; Score 548.2; DB 4; Length 787;
 Best Local Similarity 86.9%; Pred. No. 5,8e-112;
 Matches 649; Conservative 0; Mismatches 93; Indels 5; Gaps 4;

QY 687 CTCACCTGGTGGGCGGCGGAGTCTCTCTTCTAGGCGGACATTAATGTCGCTTGA 746
 Db 747 CTCGAGCTGGGCGGCGGCGGAGTCTCTCTTCTAGGCGGACATTAATGTCGCTTGA 689
 QY 747 CGGCTCATTAATTCGCGCATTCCTCAAAATCTCAATCTGATGACCTTTCC 806
 Db 688 CGGCTCATTAATTCGCGCATTCCTCAAAATCTCAATCTGATGACCTTTCC 629
 QY 807 TACTTATAGCTTAAGCATGCTGAGGCGCATGACGCGGCGGCTGTCATCTG 866
 Db 628 TACTTATAGCTTAAGCATGCTGAGGCGCATGACGCGGCGGCTGTCATCTG 569
 QY 867 TGGGCGCATCTGTACACATGCGCGGCGGCGGAGTACCTGTATCTGATGTCCTG 926
 Db 568 TGGGCGCATCTGTACACATGCGCGGCGGCGGAGTACCTGTATCTGATGTCCTG 511
 QY 927 CTCGCGGCGGCGGCGGCGGAGTATCTGAGATGATGTTCTGTCGCTTGA 986
 Db 510 CTCGCGGCGGCGGCGGCGGAGTATCTGAGATGATGTTCTGTCGCTTGA 451
 QY 987 AGTGTGCTATCTGTTGTTGTTGTTGTTGTTGTTGTTGTTGTTGTTGTTGTT 1046
 Db 450 AGTGTGCTATCTGTTGTTGTTGTTGTTGTTGTTGTTGTTGTTGTTGTTGTT 392
 QY 1047 TTTTATGTTGTTGTTGTTGTTGTTGTTGTTGTTGTTGTTGTTGTTGTTGTTGTTGTT 1106
 Db 391 TTTTATGTTGTTGTTGTTGTTGTTGTTGTTGTTGTTGTTGTTGTTGTTGTTGTTGTT 332
 QY 1107 TCCCGAAGATGCGGCTGACAGGCTGTACATGACATCTCTCTCAAGTCTGTTTC 1166
 Db 331 TCCCGAAGATGCGGCTGACAGGCTGTACATGACATCTCTCTCAAGTCTGTTTC 273
 QY 1167 CTCCTCTGTCGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 1226
 Db 272 CTCCTCTGTCGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 213
 QY 1227 TGGAAAGCTTAATTTGTCATGTCATCTAGTTTCAATTTTCTGTCGCTCTTAACAGC 1286
 Db 212 TGGAAAGCTTAATTTGTCATGTCATCTAGTTTCAATTTTCTGTCGCTCTTAACAGC 153
 QY 1287 AGTGCACCCCATCATTAATCTTCTGTCGCTGCTTGAAGGAGCTCAAAATAGGAG 1346
 Db 152 AGTGCACCCCATCATTAATCTTCTGTCGCTGCTTGAAGGAGCTCAAAATAGGAG 93
 QY 1347 AACCTGAAGCTGTTCTTCCAGAGGCTTCCAGAGACAGGCTGAGTGAAGGTTGGA 1406
 Db 92 AACCTGAAGCTGTTCTTCCAGAGGCTTCCAGAGACAGGCTGAGTGAAGGTTGGA 33
 QY 1407 GGGTGGCTTCTCAGAGAAACCTGGAG 1433
 Db 32 TGGTGTATCTCAGAGAAACCTGGAG 6

RESULT 10
 CN843258 825 bp mRNA linear EST 02-JUN-2004
 LOCUS CN843258
 DEFINITION AGENCOURT_15669738 NIH MGC_145 Homo sapiens cDNA clone
 INAME:30706610 5', mRNA sequence.
 ACCESSION CN843258
 VERSION CN843258.1 GI:47948913

KEYWORDS EST.
SOURCE Homo sapiens (human)
ORGANISM Homo sapiens
REFERENCE NIH-MGC http://mgi.nci.nih.gov/
AUTHORS National Institutes of Health, Mammalian Gene Collection (MGC)
TITLE Unpublished (1999)
JOURNAL Contact: Daniela S. Gerhard, Ph.D.
COMMENT Office of Cancer Genomics
National Cancer Institute / NIH
Bldg. 31 Rm10A07 Bethesda, MD 20892
Email: cga@dbp-1@mail.nih.gov
Tissue Procurement: GPCR Consortium
CDNA Library Preparation: GPCR Consortium
DNA Sequencing by: Agencourt Bioscience Corporation
Clone distribution: MGC clone distribution information can be found through the I.M.A.G.E. Consortium/LLNL at:
http://image.llnl.gov
Plate: IRB14 row: e column: 03
High quality sequence start: 21
High quality sequence stop: 561.
Location/Qualifiers

FEATURES
source
1..825
/organism="Homo sapiens"
/mol_type="mRNA"
/db_xref="taxon:9606"
/clone="IMAGE:30706610"
/tissue_type="mixed"
/lab_host="DH10B"
/note="Vector: pCDNA3.1; Site 1: varies by clone; Site 2: varies by clone; ORFs were PCR-amplified and cloned into pCDNA3.1 by the GPCR Consortium. Cloning sites vary by clone and include the following: 5'-EcoRV-XbaI/XhoI-3', 5'-BcoRV-XbaI/NotI-3', EcoRV (TA cloned, non-directional). For information about which gene each clones represents, please visit our anonymous ftp site at
ftp://image.llnl.gov/image/rearrayed_plates/IRB1.presv.dat
a Note: this is a NIH_MGC library."

ORIGIN
Query Match 20.8%; Score 544.6; DB 7; Length 825;
Best Local Similarity 89.5%; Pred. No. 5.1e-131;
Matches 620; Conservative 0; Mismatches 69; Indels 4; Gaps 3;

QY 761 CCGGCAATCCATCTCCAAATCT--CAGTCTGTGATGACCTTCCCTACTTTATAGGC 818
DB 725 CCCCCATCCCATCTCTAATAATCTTATCCCTGTGATGATGTTTCTACCTGCGAGGC 666
QY 819 CTAGGACATGAGGGGATGACGACGAGGGGCTGCTCATCTGTGGCCGATGCG 878
DB 665 CTGAGCTTTTGAAGCGGTGAGACCGAGCGCTGCTGCGCTGGGCCATCTGG 606
QY 879 TACCACTGCGCG-CGCCCCAGATACCTGTCATCGGTGATGTTGCTGCTGCGCCCT 937
DB 605 TACGCGTCCACGCGCCGACACACCTGTGACGGGTGGTGTGCTGCTGCGCCCT 546
QY 938 GTCCCTGCTGGAGATCTCGAGTGTGATGTTGTCGATCTTCTGTTAGTGTGCTGA 997
DB 545 GTCCCTGCTGGAGATCTCGAGTGTGATGTTAGTGTGCTTCTGTTCAAGTGTGCTGA 486
QY 998 TTTCTGTTGTTGTTGTTGTTGTTGTTGTTGTTGTTGTTGTTGTTGTTGTTGTTGTT 1057
DB 485 TTTCTGTTGTTGTTGTTGTTGTTGTTGTTGTTGTTGTTGTTGTTGTTGTTGTTGTT 426
QY 1058 GTTCTCTGTGGGTTCAGCTGTGCTGTGCTGTGCTGTGCTGTGCTGTGCTGTGCTGTG 1117
DB 425 GTTCTCTGTGGGTTCAGCTGTGCTGTGCTGTGCTGTGCTGTGCTGTGCTGTGCTGTG 366
QY 1118 GCGGCTGACGAGGCTGTGATGATGATCTCTCTCAAGTGTGTTCTCTCTCTGCG 1177

DB 365 ACCGCTGACTGAGGCTGTGATGATGATCTCTCTCAAGTGTGTTCTCTCTCTGCG 306
QY 1178 CTTGCGCTTTTGGCATTCAGTGGGCTGTGTTCCAGATTCACCTGATTTGAAAGCTT 1237
DB 305 CTTGCGCTTTTGGCATTCAGTGGGCTGTGTTCCATTTTATGATTCAGCTGACGAGAACTT 246
QY 1238 ATTGTCATGTGATCTAGTTCATTTTCTGTGCTGCTTTTAAACAGAGTGCACACC 1297
DB 245 ATTGTCATGTGATCTAGTTCATTTTCTGTGCTGCTTTTAAACAGAGTGCACACC 186
QY 1298 CATCATTTACTTCTTCTGCGGCTCCTTTAGGACGTCGCAAAATGAGAGAACCTGAAGCT 1357
DB 185 CATCATTTACTTCTTCTGCGGCTCCTTTAGGACGTCGCAAAATGAGAGAACCTGAAGCT 126
QY 1358 GTTCTCCAGAGGCTCTGACAGACGCTG-AGTGTGATGAGTGTGAGGCTGCTTC 1416
DB 125 GTTCTCCAGAGGCTCTGACAGACGCTGCTGACGATGATGAGTGTGAGGCTGCTTC 66
QY 1417 CTCAGAAATCCCTGAGCTGTGCGAAGCAGAT 1449
DB 65 CTCAGAAATCCCTGAGCTGTGCGAAGCAGAT 33

RESULT 11
CN839125 871 bp mRNA linear EST 02-JUN-2004
LOCUS CN839125
DEFINITION AGENCOURT 15864180 NIH_MGC 145 Homo sapiens CDNA clone
IMAGE:30706609 3', mRNA sequence.
ACCESSION CN839125
VERSION CN839125.1 GI:47944780
KEYWORDS EST.
SOURCE Homo sapiens
ORGANISM Homo sapiens (human)
REFERENCE Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.
1 (bases 1 to 871)
NIH-MGC http://mgi.nci.nih.gov/
AUTHORS National Institutes of Health, Mammalian Gene Collection (MGC)
TITLE Unpublished (1999)
JOURNAL Contact: Daniela S. Gerhard, Ph.D.
COMMENT Office of Cancer Genomics
National Cancer Institute / NIH
Bldg. 31 Rm10A07 Bethesda, MD 20892
Email: cga@dbp-1@mail.nih.gov
Tissue Procurement: GPCR Consortium
CDNA Library Preparation: GPCR Consortium
DNA Sequencing by: Agencourt Bioscience Corporation
Clone distribution: MGC clone distribution information can be found through the I.M.A.G.E. Consortium/LLNL at:
http://image.llnl.gov
Plate: IRB14 row: e column: 02
High quality sequence stop: 701.
Location/Qualifiers

FEATURES
source
1..871
/organism="Homo sapiens"
/mol_type="mRNA"
/db_xref="taxon:9606"
/clone="IMAGE:30706609"
/tissue_type="mixed"
/lab_host="DH10B"
/note="Vector: pCDNA3.1; Site 1: varies by clone; Site 2: varies by clone; ORFs were PCR-amplified and cloned into pCDNA3.1 by the GPCR Consortium. Cloning sites vary by clone and include the following: 5'-EcoRV-XbaI/XhoI-3', 5'-BcoRV-XbaI/NotI-3', EcoRV (TA cloned, non-directional). For information about which gene each clones represents, please visit our anonymous ftp site at
ftp://image.llnl.gov/image/rearrayed_plates/IRB1.presv.dat
a Note: this is a NIH_MGC library."

ORIGIN

Db 88 CCATGAGGAGAGCCTCTGCGCTGTGACGAGCGGACTTTGAGACACCACTGTCTTG 29
Qy 1511 C 1511
Db 28 C 28

RESULT 13
BG210740/c 871 bp mRNA linear EST 21-APR-2001
LOCUS BG210740
DEFINITION R630287 Atherys RAGE Library Homo sapiens cDNA, mRNA sequence.
ACCESSION BG210740
VERSION BG210740.1 GI:33732427
KEYWORDS EST.
SOURCE Homo sapiens (human)
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Homidae; Homo.
REFERENCE 1 (bases 1 to 871)
Harrington, J.J., Sherf, B., Rundlett, S., Jackson, P.D., Perry, R.,
Cain, S., Leventhal, C., Thornton, M., Ramchandran, R.,
Wilttington, J., Lerner, L., Costanzo, D., McEligott, K., Booser, S.,
Mays, R., Smith, E., Veloso, N., Kliska, A., Hess, J., Colgren, K., Lo, K.,
Offenbacher, J., Danzig, J. and Ducar, M.
Creation of genome-wide protein expression libraries using random
activation of gene expression
Nat. Biotechnol. 19 (5), 440-445 (2001)
JOURNAL 21227151
MEDLINE 11329013
PUBMED
COMMENT Contact: Scott J. Cain
Atherys, Inc.
3201 Carnegie Ave, Cleveland, OH 44115, USA
Tel: 216 431 9900
Fax: 216 361 9596
Email: scain@atherys.com
High quality sequence stop: 456.
Location/Qualifiers
1. 871
/organism="Homo sapiens"
/mol_type="mRNA"
/db_xref="taxon:9606"
/cell_line="HT1080"
/clone_lib="Atherys RAGE Library"
/note="See 'Creation of Genome-wide Protein Expression
Libraries using Random Activation of Gene Expression',
Nature Biotechnology, in press. Note that even though the
cell type indicated is HT1080, since a random activation
method was used, these sequence tags are not necessarily
expressed in HT1080 under normal circumstances."

ORIGIN
Query Match 18.6%; Score 486.4; DB 4; Length 871;
Best Local Similarity 90.9%; Pred. No. 9.3e-116;
Matches 528; Conservative 0; Mismatches 52; Indels 1; Gaps 1;

Qy 845 CGAGCGCTGCTGTCATCTGTGGCCCATCTGTACACATGCCCGCCCGACATACCT 904
Db 580 CGAACCCCGCTGTCCTGTCCTGT-GCCCATCTGTGTCCTGTACACATGCCCGCCCGACATACCT 522
Qy 905 GTCAATCGATATGTGTCTGCTGCGCCCTGTCCTGCTGCTGCGAGATATCTGAGTG 964
Db 521 GTCAAGCGATGTGTGTCTGCTGCTGCGCCCTGTCCTGCTGCGAGATATCTGAGTG 462
Qy 965 GATGTCGTGACTTCTGTTTATGATGCTGATCTGTTTGTGTAACGTGAGATT 1024
Db 461 GATGTTATGTGCTTCTGTTTATGATGCTGATCTGTTTGTGTAACGTGAGATT 402
Qy 1025 CATTACATCGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 1084
Db 401 CATACAGTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 342
Qy 1085 GCTGCTGAGATTCTCTGTGATCCCGAAGATGCCGCTGACGAGCTGTACGTACCT 1144
Db 1144

Db 341 GCTGATCAGGATTTCTGTGTGATCCCGAAGATCCGCTGACACAGCTGTACGTGACAT 282
Qy 1145 CTTCTCAGAGTGTGCTTCTCTCTGTGAGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 1204
Db 281 CTTCTCAGAGTGTGCTTCTCTCTGTGAGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 222
Qy 1205 GTTTTCCAGATCCAGCTGATTTGAAAGCTTTATTTTGTGATGATCTAGTTTCAT 1264
Db 221 ATTTTATGATTCACAGCTGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 162
Qy 1265 TTTCTGTCGCTCTTAAACAGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 1324
Db 161 TTTCTGTCGCTCTTAAACAGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 102
Qy 1325 TAGGAGCGCTCAAAATAGCGAAGCTGAGCTGCTTCTCAGAGGCTGCGAGAC 1384
Db 101 TAGGAGCGCTCAAAATAGCGAAGCTGAGCTGCTTCTCAGAGGCTGCGAGAGC 42
Qy 1385 GCTGAGGTGATGAGGTGAGGCTGCTTCTCAGAGAA 1425
Db 41 GCTGAGGTGATGAGGTGAGGCTGCTTCTCAGAGAA 1

RESULT 14
CD051096 924 bp mRNA linear EST 09-MAY-2003
LOCUS CD051096
DEFINITION AGENCOURT_13981113 NIH_MGC_173 Homo sapiens cDNA 5', mRNA sequence.
ACCESSION CD051096
VERSION CD051096.1 GI:30488660
KEYWORDS EST.
SOURCE Homo sapiens (human)
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Homidae; Homo.
REFERENCE 1 (bases 1 to 924)
NIH-MGC http://mgi.nci.nih.gov/
National Institutes of Health, Mammalian Gene Collection (MGC)
Unpublished (1999)
COMMENT Contact: Robert Strausberg, Ph.D.
Email: cga@bcb-rcmail.nih.gov
Tissue Procurement: Dr. Jamie Thompson, University of WI
cDNA Library Preparation: Gina Zastrow-Hayes
cDNA Library Arrayed by: The I.M.A.G.E. Consortium (LNL)
DNA Sequencing by: Agencourt Bioscience Corporation
Clone distribution: MGC clone distribution information can be
found through the I.M.A.G.E. Consortium/LNL at:
http://image.llnl.gov
Plate: NDKM39 row: e column: 24
High quality sequence start: 3
High quality sequence stop: 741.
Location/Qualifiers
1. 924
/organism="Homo sapiens"
/mol_type="mRNA"
/db_xref="taxon:9606"
/tissue_type="embryonic trophoblasts, made from WA01 stem
cells"
/clone_lib="NIH MGC 173"
/note="Vector: pDONR201, Site 1: attP2, Site 2: attP1;
LIBR PRIMING - Oligo dT; METHOD - full-length enriched;
LIBR PROVIDER - Stratfield"

ORIGIN
Query Match 18.5%; Score 483.4; DB 6; Length 924;
Best Local Similarity 87.5%; Pred. No. 5.8e-115;
Matches 552; Conservative 0; Mismatches 76; Indels 3; Gaps 2;

Qy 453 AGCAATCTCTAGGCTGATGAGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 512
Db 230 AGAATAATGAGGCTGATGAGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 289
Qy 513 TTGGGTACAGATGACCAATCAACGAGCTGAGAGACTCTTGTCTACAGAGACC 572
Db 572

```

Db 290 TTGAGACGAGAACTGACACCAATCAACGAGAGAGACTCTTTGCTCAAGAGAGACC 349
Qy 573 CTGAGCTTACGCGGGCTGAGCTGATGTTTCCCTTGTCCGCTGACAGAGAAACGGGTT 632
Db 350 TTGAGCTTACGCGGGCTGAGCTGATGTTTCCCTTGTCCGCTGACAGAGAAACGGGTT 409
Qy 633 GTGCTGCTGCTGCTGGGCTGCGGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 692
Db 410 GTGCTGCTGCTGCTGGGCTGCGGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 469
Qy 693 CTGCTGCGGCGGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 752
Db 470 TTGCGCGGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 529
Qy 753 ATGCAATATCCGCGCATCCCATCTTCAAAATCTCATGCTCTGATGAGCTTTCCCTACTT 812
Db 530 ATGCAATATCCGCGCATCCCATCTTCAAAATCTCATGCTCTGATGAGCTTTCCCTACTT 589
Qy 813 ATGAGGCTTAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 872
Db 590 GCAAGGCTTAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 649
Qy 873 ATCTGATACCACTGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCG 931
Db 650 ATCTGATACCACTGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCG 709
Qy 932 GGGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 991
Db 710 GGGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 769
Qy 992 TGCTGATCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 1051
Db 770 TGCTGATCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 829
Qy 1052 - -ATGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 1080
Db 830 TATGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 860

```

```

RESULT 15
CN837763/c 1009 bp mRNA linear EST 02-JUN-2004
LOCUS AGENOUR1.15864212 NIH MGC_145 Homo sapiens cDNA clone
DEFINITION IMAGE:30706611 3', mRNA sequence.
ACCESSION CN837763
VERSION CN837763
KEYWORDS GI:47943418
SOURCE Homo sapiens (human)
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.
REFERENCE 1 (bases 1 to 1009)
AUTHORS NIH-MGC http://mgi.nci.nih.gov/
TITLES National Institutes of Health, Mammalian Gene Collection (MGC)
JOURNAL Unpublished (1999)
COMMENT Contact: Daniela S. Gerhard, Ph.D.
Office of Cancer Genomics
National Cancer Institute / NIH
Bldg. 31 Rm10A07 Bethesda, MD 20892
Email: gcrabs-remail.nih.gov
Tissue Procurement: GPCR Consortium
cDNA Library Preparation: GPCR Consortium
DNA Library Arrayed by: The I.M.A.G.E. Consortium (LNL)
DNA Sequencing by: Agencourt Bioscience Corporation
Clone distribution: MGC clone distribution information can be
found through the I.M.A.G.E. Consortium/LNL at:
http://image.llnl.gov
Plate: IRB14 row: e column: 04
High quality sequence stop: 464.
Location/Qualifiers
1..1009
/organism="Homo sapiens"

```

```

FEATURES
Source
1..1009
/organism="Homo sapiens"

```

```

/mol_type="mRNA"
/db_xref="taxon:9606"
/clone="IMAGE:30706611"
/issue_type="mixed"
/lab_host="DH10B"
/clone_1ib="NIH MGC 145"
/notes="Vector: pCDNA3.1; Site 1: varies by clone; Site 2:
varies by clone; ORFs were PCR-amplified and cloned into
pCDNA3.1 by the GPCR Consortium. Cloning sites vary by
clone and include the following: 5'-EcoRV-XbaI/NotI-3',
5'-EcoRV-XbaI/NotI-3', EcoRV (TA cloned, non-directional).
For information about which gene each clones represents,
please visit our anonymous ftp site at
ftp://image.llnl.gov/image/rearrayed_plates/IRB1.presv.dat
a Note: this is a NIH_MGC Library."

```

```

Query Match 17.9%; Score 468.8; DB 7; Length 1009;
Best Local Similarity 88.4%; Pred. No. 4e-111;
Matches 509; Conservative 0; Mismatches 67; Indels 0; Gaps 0;

```

```

Qy 888 CGCGCGCGCGAGATACCTGTCATGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 947
Db 611 CGCGAGACAGACCGCTATCAAAAGATGATGATGATGATGATGATGATGATGATGATGAT 552
Qy 948 CGGATATCCCTGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 1007
Db 551 CGGAGATATCCCTGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 492
Qy 1008 TGTGAACGTCAGATTTCAATTAATCGGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 1067
Db 491 TGTGAACGTCAGATTTCAATTAATCGGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 432
Qy 1068 GGGTCAAGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 1127
Db 431 GGGTCAAGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 372
Qy 1128 AGGCTGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 1187
Db 371 AGGCTGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 312
Qy 1188 GGCATTAGTGAGGCGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 1247
Db 311 GGCATTAGTGAGGCGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 252
Qy 1248 GTGCACTGATTTTCATTTTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 1307
Db 251 GTTCACTGATTTTCATTTTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 192
Qy 1308 TTCTGCTGAGGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 1367
Db 191 TTCTGCTGAGGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 132
Qy 1368 AGGAGCTTGAAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 1427
Db 131 AGGAGCTTGAAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 72
Qy 1428 CTGAGAGCTGCTGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 1463
Db 71 CTGAGAGCTGCTGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 36

```

```

Search completed: October 29, 2004, 16:07:28
Job time : 7925 secs

```

This Page Blank (uspto)

GenCore version 5.1.6
Copyright (c) 1993 - 2004 CompuGen Ltd.

OM nucleic - nucleic search, using sw model

Run on: October 29, 2004, 16:11.13 : Search time 10838 Seconds
(without alignments)
11423.181 Million cell updates/sec

Title: US-09-867-570-1
Perfect score: 2618
Sequence: 1 aacaaatgcgcgaattcgcg.....gaaaaaaaaaaaaaaaaa 2618

Scoring table: OLIGO_NUC
Gapop 60.0 , Gapext 60.0

Searched: 4526729 seqs, 23644849745 residues

Word size : 20

Total number of hits satisfying chosen parameters: 105545

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Listing first 45 summaries

Database :

GenBank:.*
1: gb_ba:.*
2: gb_hgt:.*
3: gb_in:.*
4: gb_cm:.*
5: gb_ov:.*
6: gb_pat:.*
7: gb_ph:.*
8: gb_pl:.*
9: gb_pr:.*
10: gb_ro:.*
11: gb_str:.*
12: gb_sy:.*
13: gb_un:.*
14: gb_vi:.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	2618	100.0	2618	6 AX429465	AX429465 Sequence
2	1830	69.9	91923	9 AC108007	AC108007 Homo sapi
3	1830	69.9	112939	9 AC090099	AC090099 Homo sapi
4	1120	42.8	192618	2 AC020568	AC020568 Homo sapi
5	1101	42.1	8622	6 AX429467	AX429467 Sequence
6	1100	42.0	1697	9 BC067292	BC067292 Homo sapi
7	1093	41.7	1369	6 AX647081	AX647081 Sequence
8	1088	41.6	1400	6 AX299158	AX299158 Sequence
9	1088	41.6	1400	6 AX549388	AX549388 Sequence
10	1088	41.6	1400	9 AY042215	AY042215 Homo sapi
11	969	37.0	969	6 E43450	E43450 Novel prote
12	969	37.0	969	6 AX498196	AX498196 Sequence
13	918	35.1	969	6 E43451	E43451 Novel prote
14	916	35.0	969	6 AX923125	AX923125 Sequence
15	805	30.7	969	6 AF474987	AF474987 Homo sapi
16	700	26.7	912	6 AX657514	AX657514 Sequence
17	554	21.2	969	6 AR477200	AR477200 Sequence
18	554	21.2	969	6 BD103572	BD103572 Novel G P
19	554	21.2	969	9 AF474988	AF474988 Homo sapi

20	546	20.9	969	6 AR477199	AR477199 Sequence
21	546	20.9	969	6 BD103571	BD103571 Novel G P
22	386	14.7	822	6 CQ714760	CQ714760 Sequence
23	313	12.0	156839	9 AC107948	AC107948 Homo sapi
24	299	11.4	954	6 AX657472	AX657472 Sequence
25	275	10.5	275	6 AR269767	AR269767 Sequence
26	195	7.4	67908	2 AC103998	AC103998 Homo sapi
27	179	6.8	969	6 AR477201	AR477201 Sequence
28	179	6.8	969	6 BD103573	BD103573 Novel G P
29	179	6.8	969	9 AF474989	AF474989 Homo sapi
30	120	4.6	964	6 AX921833	AX921833 Sequence
31	120	4.6	969	6 AR477202	AR477202 Sequence
32	120	4.6	969	6 AX148178	AX148178 Sequence
33	120	4.6	969	6 AX268489	AX268489 Sequence
34	120	4.6	969	6 AX375428	AX375428 Sequence
35	120	4.6	969	6 AX671113	AX671113 Sequence
36	120	4.6	969	6 BD103574	BD103574 Novel G P
37	120	4.6	969	6 BD144297	BD144297 Novel G-P
38	120	4.6	969	9 AB083628	AB083628 Homo sapi
39	120	4.6	969	9 AF474990	AF474990 Homo sapi
40	120	4.6	997	6 AX921831	AX921831 Sequence
41	120	4.6	1116	6 BD095749	BD095749 Novel gua
42	120	4.6	1369	6 AX646705	AX646705 Sequence
43	120	4.6	1369	9 AB065846	AB065846 Homo sapi
44	120	4.6	2040	6 CQ815414	CQ815414 Sequence
45	120	4.6	2040	6 AX299143	AX299143 Sequence

ALIGNMENTS

RESULT 1	AX429465	Sequence 1 from Patent WO0234914.	2618 bp	DNA	linear	PAT 21-JUN-2002
LOCUS	AX429465					
DEFINITION	Sequence 1 from Patent WO0234914.					
ACCESSION	AX429465					
VERSION	AX429465.1	GI:21540763				
KEYWORDS						
SOURCE						
ORGANISM	Homo sapiens (human)					
REFERENCE	Zhao, Q., Beasley, B.M. and Wei, M.H.					
TITLE	Isolated human g-protein coupled receptors, nucleic acid molecules encoding human gpcr proteins, and uses thereof					
JOURNAL	Patent: WO 0234914-A 1 02-MAY-2002;					
PE CORP NY (US)						
FEATURES	location/Qualifiers					
SOURCE	1..2618					
ORIGIN	/organism="Homo sapiens"					
	/mol_type="unassigned DNA"					
	/db_xref="taxon:9606"					

Query Match	100.0%	Score 2618;	DB 6;	Length 2618;
Best Local Similarity	100.0%;	Pred. No. 0;		
Matches 2618;	Conservative 0;	Mismatches 0;	Indels 0;	Gaps 0;
QY	1	AACAATGCGCGAATTCGACAGATGAATCTAGTTGTTAAAGCGTGTGACACT	60	
DB	1	AACAATGCGCGAATTCGACAGATGAATCTAGTTGTTAAAGCGTGTGACACT	60	
QY	61	CT	120	
DB	61	CT	120	
QY	121	CACGAGATTGGAAGCTTCTCTGAGCGCTCCCGAAGCAAGCTGTATGCTTCTGTA	180	
DB	121	CACGAGATTGGAAGCTTCTCTGAGCGCTCCCGAAGCAAGCTGTATGCTTCTGTA	180	
QY	181	CAGTCTTAGAGCTATTAGCAGTTAAACCATTTCTTCAATAATTTCCAGTCTGAG	240	
DB	181	CAGTCTTAGAGCTATTAGCAGTTAAACCATTTCTTCAATAATTTCCAGTCTGAG	240	

Dh 181 CAGTCTGAGACGATTAAGCCAGTTAAACCAATTCCTTCAATTAATTTCCAGTCTCAGG 240
Qy 241 TATTTCTTTTGAAGATTTGAGATGAACTAAATACACAGACAGAGAGCCAGAGATGAA 300
Dh 241 TATTTCTTTTGAAGATTTGAGATGAACTAAATACACAGAGAGCCAGAGATGAA 300
Qy 301 ATCCCAAGGTGCTTTCTGCTGTCTTCCAGTCTCTGCTGTGTGTCTCCAGTGTCTCAAT 360
Dh 301 ATCCCAAGGTGCTTTCTGCTGTCTTCCAGTCTCTGCTGTGTGTCTCCAGTGTCTCAAT 360
Qy 361 TCCACACAGAAACCAAAATTAAGAAATCCCACTGATGTGTGATCAATAGAACCACTCTCT 420
Dh 361 TCCACACAGAAACCAAAATTAAGAAATCCCACTGATGTGTGATCAATAGAACCACTCTCT 420
Qy 421 TGGGATGTCAAAACAGATTAAGAGATGAAAGCAAAATCTCATGSGTCAATCAGACTG 480
Dh 421 TGGGATGTCAAAACAGATTAAGAGATGAAAGCAAAATCTCATGSGTCAATCAGACTG 480
Qy 481 GGTTCCTGAGCATGGAATTCACATCCAGTCTTGSGTACAGAACTGACACCAATCAAG 540
Dh 481 GGTTCCTGAGCATGGAATTCACATCCAGTCTTGSGTACAGAACTGACACCAATCAAG 540
Qy 541 GACGTGAGAGACCTCTGCTACAGAGAGACCTGAGCTTCAAGGGGCTGAGAGTGCATG 600
Dh 541 GACGTGAGAGACCTCTGCTACAGAGAGACCTGAGCTTCAAGGGGCTGAGAGTGCATG 600
Qy 601 TTTCCCTGTGCGCTGACAGAGAAACGCGTGTGTGCTCTGAGCTCTGAGGCTGCCAGATC 660
Dh 601 TTTCCCTGTGCGCTGACAGAGAAACGCGTGTGTGCTCTGAGGCTGCCAGATC 660
Qy 661 GCAGAGACGTGTGTCTCACTACATCTCAACCTGTGTGCGGCGCACTTCTCTTCTTA 720
Dh 661 GCAGAGACGTGTGTCTCACTACATCTCAACCTGTGTGCGGCGCACTTCTCTTCTTA 720
Qy 721 GCGGCAACATTATAGTTGCGCGTTACGCGTCAATCAATCCGCGCACTCCATCTCCAAA 780
Dh 721 GCGGCAACATTATAGTTGCGCGTTACGCGTCAATCAATCCGCGCACTCCATCTCCAAA 780
Qy 781 TCTCTAGTCTGTGATGACCTTTCCCTAATTATAGGCGCTAAGCATGCTGAGGCGCATCA 840
Dh 781 TCTCTAGTCTGTGATGACCTTTCCCTAATTATAGGCGCTAAGCATGCTGAGGCGCATCA 840
Qy 841 GCACCGAGCGCTGCTGTCTCACTCTGTGCGCATCTGTGATCACTGCGCGCGCCAGAT 900
Dh 841 GCACCGAGCGCTGCTGTCTCACTCTGTGCGCATCTGTGATCACTGCGCGCGCCAGAT 900
Qy 901 ACCGTGATCGGTGATGTGTCTGCTGTGAGGCGCTGCGCGCGCGCGAGATCTGAG 960
Dh 901 ACCGTGATCGGTGATGTGTCTGCTGTGAGGCGCTGCGCGCGCGCGAGATCTGAG 960
Qy 961 AGTGAATGTCTGTGATCTCTGTTAGTGTGTGATCTGTGTGTGTGAAACGTCTAG 1020
Dh 961 AGTGAATGTCTGTGATCTCTGTTAGTGTGTGATCTGTGTGTGTGAAACGTCTAG 1020
Qy 1021 ATTTTCAATTAACAATCGCGTGTGTTTATATGTGTGTCTGTGTGTGTCAAGCTG 1080
Dh 1021 ATTTTCAATTAACAATCGCGTGTGTTTATATGTGTGTCTGTGTGTGTCAAGCTG 1080
Qy 1081 TCTGCTGTGTGATGATCTGTGTGATCCGGAAGATGCGGCTGACCAAGGCTGTAGCTGA 1140
Dh 1081 TCTGCTGTGTGATGATCTGTGTGATCCGGAAGATGCGGCTGACCAAGGCTGTAGCTGA 1140
Qy 1141 CCATCTCTCTCAAGATGTGTCTCTCTCTCTGTGTGCTGTGCTGTGTGCAATTCAGTGG 1200
Dh 1141 CCATCTCTCTCAAGATGTGTCTCTCTCTCTGTGTGCTGTGCTGTGTGCAATTCAGTGG 1200
Qy 1201 CCGTGTGTGTCAAGATCACTGTGTGATGAAAGTCTTATTTGTGTGTGTGTGTGTGTGT 1260
Dh 1201 CCGTGTGTGTCAAGATCACTGTGTGATGAAAGTCTTATTTGTGTGTGTGTGTGTGTGT 1260
Qy 1261 CCATTTTCTGTGCGCTCTTAAACAGAGTGCACCCCATCAATTAATCTTCTGTGTGTGT 1320
Dh 1261 CCATTTTCTGTGCGCTCTTAAACAGAGTGCACCCCATCAATTAATCTTCTGTGTGTGT 1320

Qy 1321 CCTTTAGGCAAGGTCAAAATAGCAGAACTGAACTGTCTCTCAGAGGCTCTGCAGG 1380
Dh 1321 CCTTTAGGCAAGGTCAAAATAGCAGAACTGAACTGTCTCTCAGAGGCTCTGCAGG 1380
Qy 1381 ACAAGCTGTGTGATGAAAGTGTGAGGCTGTCTCCAGAGAAACCTGTGAGCTGTG 1440
Dh 1381 ACAAGCTGTGTGATGAAAGTGTGAGGCTGTCTCCAGAGAAACCTGTGAGCTGTG 1440
Qy 1441 GAAGCAGATTTGAGCAGTGAAGAGAAACCTGTGCGCTGTCAAGACAGACTTTGAGAGCA 1500
Dh 1441 GAAGCAGATTTGAGCAGTGAAGAGAAACCTGTGCGCTGTCAAGACAGACTTTGAGAGCA 1500
Qy 1501 TGTGCGCTGTCAAGCTTTGACATTAATGCAATTTTCTTAAGCTTCTGTCTCAGAAATG 1560
Dh 1501 TGTGCGCTGTCAAGCTTTGACATTAATGCAATTTTCTTAAGCTTCTGTCTCAGAAATG 1560
Qy 1561 TCTCAAGTGTCTCTCAAGTCTTCAAAATGATGTTAATCTTAACCTGACAGATTTGAGTTT 1620
Dh 1561 TCTCAAGTGTCTCTCAAGTCTTCAAAATGATGTTAATCTTAACCTGACAGATTTGAGTTT 1620
Qy 1621 CACCATGGAAGATTAAGTCTGACAGTACAAATGTTGGAATTTCTCTGATATTAACAT 1680
Dh 1621 CACCATGGAAGATTAAGTCTGACAGTACAAATGTTGGAATTTCTCTGATATTAACAT 1680
Qy 1681 ACAATTTCTCTGTATCTTGAACCTGAATCTTCTCTACTGAACACTTTTCTGCACTTTTC 1740
Dh 1681 ACAATTTCTCTGTATCTTGAACCTGAATCTTCTCTACTGAACACTTTTCTGCACTTTTC 1740
Qy 1741 ATTTGTAATTAAGAGTGTGTGTGCAAAACCTTAACCTCTTTTAATCTTGTGTTCTTA 1800
Dh 1741 ATTTGTAATTAAGAGTGTGTGTGCAAAACCTTAACCTCTTTTAATCTTGTGTTCTTA 1800
Qy 1801 CCTGATAGTACAAAGAGATTCCTTAATATCTGTCAAGACTATGTTCCCTGAAAA 1860
Dh 1801 CCTGATAGTACAAAGAGATTCCTTAATATCTGTCAAGACTATGTTCCCTGAAAA 1860
Qy 1861 TCAATGTTCCCTTTATGACTGAGGCAATTAAGCTGTGAGAGCTCAATCTTAAATAGT 1920
Dh 1861 TCAATGTTCCCTTTATGACTGAGGCAATTAAGCTGTGAGAGCTCAATCTTAAATAGT 1920
Qy 1921 GAGTCTGTGATCTCTAAATTTCCATGAAATCTGATTAATAGCAAAATTAATAGCTTA 1980
Dh 1921 GAGTCTGTGATCTCTAAATTTCCATGAAATCTGATTAATAGCAAAATTAATAGCTTA 1980
Qy 1981 GAGAGAGATTCCTCTCAATTAAGAGCTGTAGAAATGTTGTTATGAAATAGCCCTC 2040
Dh 1981 GAGAGAGATTCCTCTCAATTAAGAGCTGTAGAAATGTTGTTATGAAATAGCCCTC 2040
Qy 2041 CTGTCAATTTGTCAACAGATGTGTGATGCTGTGCTGTGTCTTAAGAACATCTGT 2100
Dh 2041 CTGTCAATTTGTCAACAGATGTGTGATGCTGTGCTGTGTCTTAAGAACATCTGT 2100
Qy 2101 GCGCCCTCCCTTGAGACTGTGATGCTTAATTAAGCTTCTCGGACTAATGAACTAG 2160
Dh 2101 GCGCCCTCCCTTGAGACTGTGATGCTTAATTAAGCTTCTCGGACTAATGAACTAG 2160
Qy 2161 TGAGAGGCTTAATTAATGATCCACCACTTCAATTTTGCCATTTGGAACCTCAATATG 2220
Dh 2161 TGAGAGGCTTAATTAATGATCCACCACTTCAATTTTGCCATTTGGAACCTCAATATG 2220
Qy 2221 ATTTTAAAGTGAATTAATCTTGAAGAACATTTATTAATCACTTAAGATCTTCAAGTT 2280
Dh 2221 ATTTTAAAGTGAATTAATCTTGAAGAACATTTATTAATCACTTAAGATCTTCAAGTT 2280
Qy 2281 GTAGAGAAATCTTATATCTTCCAGTGTGTGATTAATGTTCTGATTTGATGAACTTCA 2340
Dh 2281 GTAGAGAAATCTTATATCTTCCAGTGTGTGATTAATGTTCTGATTTGATGAACTTCA 2340
Qy 2341 TAGTTTATGCTGTGTTACATGAGAGCAAACTGAACATCTGACCTTTCAATGACAA 2400
Dh 2341 TAGTTTATGCTGTGTTACATGAGAGCAAACTGAACATCTGACCTTTCAATGACAA 2400

Oy		2401	TCTCAATTATGTAATCGGAATAAATACTACAGTGGACAGAATTCTGAATACATGCTGT	2460
Dd		2401	TCTCAATTATGTAATCTCGAATAAATACTACAGTTGGACAGAATTCTGAATACATGCTGT	2460
Oy		2461	GACATACATGAACTCGGAAAATATTTGTGCTGAAGAAAAATTAAGCCAGACGCCCAACAATATT	2520
Dd		2461	GACATACATGAACTCGGAAAATATTGTGCTGAAGAAAAATTAAGCCAGACGCCCAACAATATT	2520
Oy		2521	GTAA GTTC AAAT TCTTA TGAGGTATTC CAAATTAG GAAATTC TTGTA ACACAG AATAATTA AT	2580
Dd		2521	GTAA GTTC AAAT TCTTA TGAGGTATTC CAAATTAG GAAATTC TTGTA ACACAG AATAATTA AT	2580
Oy		2581	AGAGAGATCCTGCTGCATGAAAAAAAAAAAAAAAAAAAAA	2618
Dd		2581	AGAGAGATCCTGCTGCATGAAAAAAAAAAAAAAAAAAAAA	2618
RESULT 2				
ACI08007/c				
LOCUS	ACI08007	91923 bp	DNA	linear PRI 03-SEP-2002
DEFINITION	Homo sapiens chromosome 11, clone CTD-3038L12, complete sequence.			
ACCESSION	ACI08007			
VERSION	ACI08007.5	GI:22655822		
KEYWORDS	HTG.			
SOURCE	Homo sapiens (human)			
ORGANISM	Homo sapiens			
REFERENCE	Bukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Primates; Catarrhini; Hominiidae; Homo. 1 (bases 1 to 91923) Birten,B., Nusbaum,C. and Lander,E. Homo sapiens chromosome 11, clone CTD-3038L12 Unpublished 2 (bases 1 to 91923) Birten,B., Binton,L., Nusbaum,C., Lander,E., Ali,A., Allen,N.,			
AUTHORS				

Brown, A., Canarata, J., Campojano, A., Chang, J., Chazaro, B.,
 Choepel, Y., Colangelo, M., Collins, S., Collymore, A., Cook, A.,
 Cooke, P., DeKrellano, K., Dewar, K., Diaz, J.S., Dodge, S., Fato, S.,
 Ferreira, P., FitzHugh, W., Gage, D., Galagan, J., Gardyna, S.,
 Ginde, S., Gord, S., Goyette, M., Graham, L., Grand-Pierre, N.,
 Hagos, B., Horton, L., Hulme, W., Iliev, I., Johnson, R., Jones, C.,
 Kanat, A., Karatas, A., Kells, C., Lacroque, K., Lamaatras, R.,
 Landers, T., Lehoczy, J., Levine, R., Liu, G., Maclean, C.,
 Macdonald, P., Major, J., Marguis, N., Matthews, C., McCarthy, M.,
 Mcwan, P., McKernan, K., Meldrum, J., Menus, L., Mihova, T.,
 Mienga, V., Murphy, T., Naylor, J., Nguyen, C., Nicol, R., Norbu, C.,
 Norman, C.H., O'Connor, T., O'Donnell, P., O'Neill, D., Oliver, J.,
 Peterson, K., Phunhkiang, P., Pierre, N., Pollara, V., Raymond, C.,
 Retta, R., Ribbeck, M., Riley, R., Rise, C., Rogov, P., Roman, J.,
 Rosetti, M., Roy, A., Santos, R., Schauer, S., Schupbach, R., Seaman, S.,
 Severy, P., Spencer, B., Strange-Thomann, N., Stojanovic, N.,
 Straus, N., Subramanian, A., Talmas, J., TeeEaye, S., Theodore, J.,
 Topham, K., Travers, M., Travis, N., Triggillo, J., Vassiliev, H.,
 Viel, R., Vo, A., Wilson, B., Wu, X., Wyman, D., Ye, W. J., Young, G.,
 Zelnoun, J., Zemбек, L., Zimmer, A. and Zody, M.

Direct Submission
 Submitted (24-JAN-2002) Whitehead Institute/MIT Center for Genome
 Research, 320 Charles Street, Cambridge, MA 02141, USA
 3 (bases 1 to 91923)

Birten, B., Nishum, C., Lander, E., Ali, A., Allen, N., Anderson, S.,
 Barua, N., Baetien, V., Bloom, T., Boguslavskiy, L., Bouthgalter, B.,
 Canarata, J., Chang, J., Chazaro, B., Choepel, Y., Collymore, A.,
 Cook, A., Cooke, P., DeKrellano, K., Dewar, K., Diaz, J.S., Dodge, S.,
 Fato, S., Ferreira, P., Fitzgerald, M., Gage, D., Galagan, J.,
 Gardyna, S., Gord, S., Graham, L., Grand-Pierre, N., Hagos, B.,
 Horton, L., Hulme, W., Iliev, I., Johnson, R., Jones, C., Kanat, A.,
 Karatas, A., Kells, C., Landers, T., Levine, R., Lindblad-Toh, K.,
 Liu, G., Maclean, C., Macdonald, P., Major, J., Matthews, C.,
 McCarthy, M., Meldrum, J., Menus, L., Mihova, T., Mienga, V.,
 Murphy, T., Naylor, J., Nguyen, C., Nicol, R., Norbu, C., Norman, C.H.,
 O'Connor, T., O'Donnell, P., O'Neill, D., Oliver, J., Peterson, K.,
 Phunhkiang, P., Pierre, N., Raymond, C., Retta, R., Rise, C., Rogov, P.,
 Roman, J., Roy, A., Schauer, S., Schupbach, R., Seaman, S., Severy, P.,

REFERENCE

TITLE
JOURNAL

AUTHORS

Smith,C., Spencer,B., Stange-Thomann,N., Stojanovic,N., Talamaas,J.,
Tsafaye,S., Theodore,J., Topham,K., Travers,M., Vassiliev,H.,
Viel,R., Vo,A., Wilson,B., Wu,X., Wyman,D., Young,G., Zainoun,J.,
Zembek,L., Zimmer,A. and Zody,M.

Direct Submission
Submitted (20-APR-2002) Whitehead Institute/MIT Center for Genome
Research, 320 Charles Street, Cambridge, MA 02141, USA

4 (bases 1 to 91923)

Birren,B., Nussbaum,C., Lander,E., Ali,A., Allen,N., Anderson,S.,
Barna,N., Bassett,V., Bloom,T., Boguski,K.Y., Boulanger,B.,
Camarata,J., Chang,J., Chazaro,B., Choquel,Y., Collymore,A.,
Cook,A., Cooke,P., DeArillano,K., Dewar,K., Diaz,J.S., Dodge,S.,
Faro,S., Ferreira,P., FitzGerald,M., Gage,D., Galagan,J.,
Gardyna,S., Goid,S., Graham,L., Grand-Pierre,N., Hagos,B.,
Horton,L., Hulme,W., Iliev,I., Johnson,R., Jones,C., Kanat,A.,
Karatas,A., Kellis,C., Landert,T., Levine,R., Lindblad-Toh,K.,
Liu,G., Maclean,C., Macdonald,P., Major,J., Matthews,C.,
McCarthy,M., Meldrum,J., Menus,L., Mihova,T., Mleung,V.,
Murphy,T., Naylor,J., Nguyen,C., Nicol,R., Norbu,C., Norman,C.H.,
O'Connor,T., O'Donnell,P., O'Neill,D., Oliver,J., Peterson,K.,
Phunkhang,P., Pierre,N., Raymond,C., Retta,R., Rise,C., Rogov,P.,
Roman,J., Roy,A., Schauer,S., Schuppach,R., Seaman,S., Severy,P.,
Smith,C., Spencer,B., Stange-Thomann,N., Stojanovic,N., Talamaas,J.,
Tsafaye,S., Theodore,J., Topham,K., Travers,M., Vassiliev,H.,
Viel,R., Vo,A., Wilson,B., Wu,X., Wyman,D., Young,G., Zainoun,J.,
Zembek,L., Zimmer,A. and Zody,M.

Direct Submission
Submitted (03-SEP-2002) Whitehead Institute/MIT Center for Genome
Research, 320 Charles Street, Cambridge, MA 02141, USA

On Sep 3, 2002 this sequence version replaced gi:22474954.

```

Center code: WtBR
Web site: http://www-seg.wi.mit.edu
Contact: sequence_submissions@genome.wi.mit.edu
----- Project Information -----
Center project name: L24596
Center clone name: 3038_L12
-----

Only the first 91.9 kb of this clone are being submitted.
The remainder overlaps accession number AC055860 [WICGR project
L9785].

FEATURES
        source
            1. .91923
                /organism="Homo sapiens"
                /mol_type="genomic DNA"
                /db_xref="taxon:9606"
                /chromosome="11"
                /map="11"
                /clone="CTD-3038L12"
                /clone_11b="CTD2 Human BAC"
            25. .76
                /rpt_family="L2"
                /complement(251. .899)
                /rpt_family="L2"
                900. .1001
                /rpt_family="L1MD2"
                1082. .1655
                /rpt_family="L1MD2"
                1671. .1850
                /rpt_family="L1MD3"
                complement(1658. .1917)
                /rpt_family="L2"
                1918. .2273
                /rpt_family="THE1B"
                complement(2274. .2327)
                /rpt_family="L2"
                4505. .4713
                /rpt_family="L2"

```

```
repeat_region 5536..5795 /rpe_family="AluSc"
repeat_region 5796..5837 /rpe_family="TAAAn"
repeat_region 5838..5875 /rpe_family="AluSc"
repeat_region 5876..5917 /rpe_family="L1MB4"
repeat_region 5918..5959 /rpe_family="MIR"
repeat_region 5960..5999 /rpe_family="L12"
repeat_region 6000..6039 /rpe_family="MIR3"
repeat_region 6040..6079 /rpe_family="MER2"
repeat_region 6080..6119 /rpe_family="L1MB3"
repeat_region 6120..6159 /rpe_family="MSTa"
repeat_region 6160..6199 /rpe_family="L1MB3"
repeat_region 6200..6239 /rpe_family="L1MB1"
repeat_region 6240..6279 /rpe_family="L1MB1"
repeat_region 6280..6319 /rpe_family="L1MB1"
repeat_region 6320..6359 /rpe_family="L1MB1"
repeat_region 6360..6399 /rpe_family="L1MB1"
repeat_region 6400..6439 /rpe_family="L1MB1"
repeat_region 6440..6479 /rpe_family="L1MB1"
repeat_region 6480..6519 /rpe_family="L1MB1"
repeat_region 6520..6559 /rpe_family="L1MB1"
repeat_region 6560..6599 /rpe_family="L1MB1"
repeat_region 6600..6639 /rpe_family="L1MB1"
repeat_region 6640..6679 /rpe_family="L1MB1"
repeat_region 6680..6719 /rpe_family="L1MB1"
repeat_region 6720..6759 /rpe_family="L1MB1"
repeat_region 6760..6799 /rpe_family="L1MB1"
repeat_region 6800..6839 /rpe_family="L1MB1"
repeat_region 6840..6879 /rpe_family="L1MB1"
repeat_region 6880..6919 /rpe_family="L1MB1"
repeat_region 6920..6959 /rpe_family="L1MB1"
repeat_region 6960..6999 /rpe_family="L1MB1"
repeat_region 7000..7039 /rpe_family="L1MB1"
repeat_region 7040..7079 /rpe_family="L1MB1"
repeat_region 7080..7119 /rpe_family="L1MB1"
repeat_region 7120..7159 /rpe_family="L1MB1"
repeat_region 7160..7199 /rpe_family="L1MB1"
repeat_region 7200..7239 /rpe_family="L1MB1"
repeat_region 7240..7279 /rpe_family="L1MB1"
repeat_region 7280..7319 /rpe_family="L1MB1"
repeat_region 7320..7359 /rpe_family="L1MB1"
repeat_region 7360..7399 /rpe_family="L1MB1"
repeat_region 7400..7439 /rpe_family="L1MB1"
repeat_region 7440..7479 /rpe_family="L1MB1"
repeat_region 7480..7519 /rpe_family="L1MB1"
repeat_region 7520..7559 /rpe_family="L1MB1"
repeat_region 7560..7599 /rpe_family="L1MB1"
repeat_region 7600..7639 /rpe_family="L1MB1"
repeat_region 7640..7679 /rpe_family="L1MB1"
repeat_region 7680..7719 /rpe_family="L1MB1"
repeat_region 7720..7759 /rpe_family="L1MB1"
repeat_region 7760..7799 /rpe_family="L1MB1"
repeat_region 7800..7839 /rpe_family="L1MB1"
repeat_region 7840..7879 /rpe_family="L1MB1"
repeat_region 7880..7919 /rpe_family="L1MB1"
repeat_region 7920..7959 /rpe_family="L1MB1"
repeat_region 7960..7999 /rpe_family="L1MB1"
repeat_region 8000..8039 /rpe_family="L1MB1"
repeat_region 8040..8079 /rpe_family="L1MB1"
repeat_region 8080..8119 /rpe_family="L1MB1"
repeat_region 8120..8159 /rpe_family="L1MB1"
repeat_region 8160..8199 /rpe_family="L1MB1"
repeat_region 8200..8239 /rpe_family="L1MB1"
repeat_region 8240..8279 /rpe_family="L1MB1"
repeat_region 8280..8319 /rpe_family="L1MB1"
repeat_region 8320..8359 /rpe_family="L1MB1"
repeat_region 8360..8399 /rpe_family="L1MB1"
repeat_region 8400..8439 /rpe_family="L1MB1"
repeat_region 8440..8479 /rpe_family="L1MB1"
repeat_region 8480..8519 /rpe_family="L1MB1"
repeat_region 8520..8559 /rpe_family="L1MB1"
repeat_region 8560..8599 /rpe_family="L1MB1"
repeat_region 8600..8639 /rpe_family="L1MB1"
repeat_region 8640..8679 /rpe_family="L1MB1"
repeat_region 8680..8719 /rpe_family="L1MB1"
repeat_region 8720..8759 /rpe_family="L1MB1"
repeat_region 8760..8799 /rpe_family="L1MB1"
repeat_region 8800..8839 /rpe_family="L1MB1"
repeat_region 8840..8879 /rpe_family="L1MB1"
repeat_region 8880..8919 /rpe_family="L1MB1"
repeat_region 8920..8959 /rpe_family="L1MB1"
repeat_region 8960..8999 /rpe_family="L1MB1"
repeat_region 9000..9039 /rpe_family="L1MB1"
repeat_region 9040..9079 /rpe_family="L1MB1"
repeat_region 9080..9119 /rpe_family="L1MB1"
repeat_region 9120..9159 /rpe_family="L1MB1"
repeat_region 9160..9199 /rpe_family="L1MB1"
repeat_region 9200..9239 /rpe_family="L1MB1"
repeat_region 9240..9279 /rpe_family="L1MB1"
repeat_region 9280..9319 /rpe_family="L1MB1"
repeat_region 9320..9359 /rpe_family="L1MB1"
repeat_region 9360..9399 /rpe_family="L1MB1"
repeat_region 9400..9439 /rpe_family="L1MB1"
repeat_region 9440..9479 /rpe_family="L1MB1"
repeat_region 9480..9519 /rpe_family="L1MB1"
repeat_region 9520..9559 /rpe_family="L1MB1"
repeat_region 9560..9599 /rpe_family="L1MB1"
repeat_region 9600..9639 /rpe_family="L1MB1"
repeat_region 9640..9679 /rpe_family="L1MB1"
repeat_region 9680..9719 /rpe_family="L1MB1"
repeat_region 9720..9759 /rpe_family="L1MB1"
repeat_region 9760..9799 /rpe_family="L1MB1"
repeat_region 9800..9839 /rpe_family="L1MB1"
repeat_region 9840..9879 /rpe_family="L1MB1"
repeat_region 9880..9919 /rpe_family="L1MB1"
repeat_region 9920..9959 /rpe_family="L1MB1"
repeat_region 9960..9999 /rpe_family="L1MB1"
```

```
Query Match 69.9%; Score 1830; DB 9; Length 91923;
Best Local Similarity 99.7%; Pred. No. 0;
Matches 2130; Conservative 0; Mismatches 6; Indels 0; Gaps 0;

466 GGGTCATAGAGCTGGGGTTTCTGAGCATTGATTCATCCAGTCTGGGTACAGAC 525
Db 9503 GGGTCATAGAGCTGGGGTTTCTGAGCATTGATTCATCCAGTCTGGGTACAGAC 9444
526 TGACACCAATCAAGAGAGTGAAGAGTCTCTTCTCAAGACAGACCTGAGCTTCAAG 585
Db 9443 TGACACCAATCAAGAGAGTGAAGAGTCTCTTCTCAAGACAGACCTGAGCTTCAAG 9384
586 GAGTCAGTGAAGTGTTCCTTGTCTGCTGAGAGAGAGAGAGAGAGAGAGAGAGAG 645
Db 9383 GAGTCAGTGAAGTGTTCCTTGTCTGCTGAGAGAGAGAGAGAGAGAGAGAGAGAG 9324
646 TGGGCTGCGGAGTGGGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 705
Db 9323 TGGGCTGCGGAGTGGGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 9284
706 ACTTCCTCTTCCTTACGCGGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 765
Db 9283 ACTTCCTCTTCCTTACGCGGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 9204
766 ATCCATCTCCAAATCTCAAGTCTGATGACCTTCTCTTATAGGCTTAAGCA 825
Db 9203 ATCCATCTCCAAATCTCAAGTCTGATGACCTTCTCTTATAGGCTTAAGCA 9144
826 TGGTCAGGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 885
Db 9143 TGGTCAGGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 9084
886 GCGCGCGCGCGAGATACCTGATCAGTGTGATGATGATGATGATGATGATGATGAT 945
Db 9083 GCGCGCGCGCGAGATACCTGATCAGTGTGATGATGATGATGATGATGATGATGAT 9024
946 TGGGAGATATCTGAGAGAGAGATGATGATGATGATGATGATGATGATGATGATG 1005
Db 9023 TGGGAGATATCTGAGAGAGAGATGATGATGATGATGATGATGATGATGATGATG 8964
1006 GGTGTGAAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 1065
Db 8963 GGTGTGAAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 8904
1066 GTGGTCAAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 1125
Db 8903 GTGGTCAAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 8844
1126 CAGGCTGATGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 1185
Db 8843 CAGGCTGATGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 8784
1186 TTGGATTCAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 1245
Db 8783 TTGGATTCAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 8724
1246 ATGTGATCTAGATTCATTTCTGCTGCTGCTTAAAGAGAGAGAGAGAGAGAGAG 1305
Db 8723 ATGTGATCTAGATTCATTTCTGCTGCTGCTTAAAGAGAGAGAGAGAGAGAGAG 8664
1306 ACTTCTGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 1365
Db 8663 ACTTCTGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 8604
1366 AGAGGAGCTGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 1425
Db 8603 AGAGGAGCTGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 8544
1426 CCTGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 1485
Db 8543 CCTGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 8484
```


QY 1486 GAACTTTGAGAGCAATGCTGCCCTGCCACCTTGACATATATATGACATTTTCTTAGCCT 1545
 DB 8483 GAACTTTGAGAGCAATGCTGCCCTGCCACCTTGACATATATATGACATTTTCTTAGCCT 8424
 QY 1546 TCTGCTCAGAAATGCTCAGTGTCCCTCAAGCTCTTGAAATGATGTTTATCTAACCT 1605
 DB 8423 TCTGCTCAGAAATGCTCAGTGTCCCTCAAGCTCTTGAAATGATGTTTATCTAACCT 8364
 QY 1606 GACAGTTGAGCTTTGACCCCAATGGAAGCAATGCTGACAGTGCATGTTTGAATCTC 1665
 DB 8363 GACAGTTGAGCTTTGACCCCAATGGAAGCAATGCTGACAGTGCATGTTTGAATCTC 8304
 QY 1666 CTGATATATACCAATATATTTCCCTGTTATCTTGACATGATCTTTCTTACGACCT 1725
 DB 8303 CTGATATATACCAATATATTTCCCTGTTATCTTGACATGATCTTTCTTACGACCT 8244
 QY 1726 TTTTCTGCACTTTTCATTTGTAATAAAGAGTGTCTGTCCAAACCTTAAACTCTTCTT 1785
 DB 8243 TTTTCTGCACTTTTCATTTGTAATAAAGAGTGTCTGTCCAAACCTTAAACTCTTCTT 8184
 QY 1786 TATATCTGTTTCCCACTGATGATGATGATGATGATGATGATGATGATGATGATGAT 1845
 DB 8183 TATATCTGTTTCCCACTGATGATGATGATGATGATGATGATGATGATGATGATGAT 8124
 QY 1846 ATGTTCCCTGAAATATGATGTTCCCTTTTATGACATGAGCAATGACATGAGTGAAGCT 1905
 DB 8123 ATGTTCCCTGAAATATGATGTTCCCTTTTATGACATGAGCAATGACATGAGTGAAGCT 8064
 QY 1906 CAATTTCTTAATTAAGTGAATGTTGCTACCTCTTAAATTCATTTGAATTTCTCAGATTAAGC 1965
 DB 8063 CAATTTCTTAATTAAGTGAATGTTGCTACCTCTTAAATTCATTTGAATTTCTCAGATTAAGC 8004
 QY 1966 AAAATATGACCTTGAAGAGATGTTCCCTTCAATTAATAAAGCTTGAATTTGTTT 2025
 DB 8003 AAAATATGACCTTGAAGAGATGTTCCCTTCAATTAATAAAGCTTGAATTTGTTT 7944
 QY 2026 ATGAATAGCCCTCTCCTGTCATTTGTCACAGATGAGTGAATGAGCTTGTGTTCTA 2085
 DB 7943 ATGAATAGCCCTCTCCTGTCATTTGTCACAGATGAGTGAATGAGCTTGTGTTCTA 7884
 QY 2086 GTAAAGCAATCTGAGCCCTTCCCTTGAGAACTGTGAATTTTATGATCTTCTT 2145
 DB 7883 GTAAAGCAATCTGAGCCCTTCCCTTGAGAACTGTGAATTTTATGATCTTCTTCTT 7824
 QY 2146 GGAATATGACCTTGAAGAGATGATGATGATGATGATGATGATGATGATGATGATGAT 2205
 DB 7823 GGAATATGACCTTGAAGAGATGATGATGATGATGATGATGATGATGATGATGATGAT 7764
 QY 2206 GAAACCTCAATATGATTTTAAATGGAATTTATCTTGAAACCAATTTATTTACCTTA 2265
 DB 7763 GAAACCTCAATATGATTTTAAATGGAATTTATCTTGAAACCAATTTATTTACCTTA 7704
 QY 2266 CAGATTTCTCAGTTGAGAGAAATCTTCAATCTTCAGGTTTGTATAAATGTTCTG 2325
 DB 7703 CAGATTTCTCAGTTGAGAGAAATCTTCAATCTTCAGGTTTGTATAAATGTTCTG 7644
 QY 2326 ATGTAATCTTCACTGATGTTTATGAGCTGTTTATGAGAGAAACCAATCTGAAACATCTG 2385
 DB 7643 ATGTAATCTTCACTGATGTTTATGAGCTGTTTATGAGAGAAACCAATCTGAAACATCTG 7584
 QY 2386 ACCTTTCATGACATCTCAATATGATGATGATGATGATGATGATGATGATGATGATGAT 2445
 DB 7583 ACCTTTCATGACATCTCAATATGATGATGATGATGATGATGATGATGATGATGATGAT 7524
 QY 2446 TCTGATATCATGCTGTCATATCATGAAACCTGGAATATGTTGCTAAGAGAAATTAAGCCAG 2505
 DB 7523 TCTGATATCATGCTGTCATATCATGAAACCTGGAATATGTTGCTAAGAGAAATTAAGCCAG 7464
 QY 2506 ACGGCAAAACATATTTGTAAGTTCATTTCTATGAGGATTCATTAATTTAGAAATTTCTTGA 2565
 DB 7463 ACGGCAAAACATATTTGTAAGTTCATTTCTATGAGGATTCATTAATTTAGAAATTTCTTGA 7404
 QY 2566 CACGAAATTAATTAAGAGATCTGCTGTCTGGA 2601

DB 7403 CACGAAATTAATTAAGAGATCTGCTGTCTGGA 7368
 RESULT 3
 AC090099/C
 LOCUS
 DEFINITION
 ACCESSION
 VERSION
 KEYWORDS
 SOURCE
 ORGANISM
 Homo sapiens
 Homo sapiens (human)
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.
 AC090099.14 GI:23396332
 HTS.
 172939 bp DNA linear PRI 01-OCT-2002
 Homo sapiens chromosome 11, clone RP11-113D6, complete sequence.
 1 (bases 1 to 172939)
 Homo sapiens chromosome 11, clone RP11-113D6
 Unpublished
 2 (bases 1 to 172939)
 REFERENCES
 AUTHORS
 Birren, B., Linton, L., Nusbaum, C., Lander, E., Allen, N., Anderson, S.,
 Barna, N., Bastien, V., Boguslavsky, L., Boukhalter, B., Brown, A.,
 Camarata, J., Campopiano, A., Choquet, Y., Colangelo, M., Collins, S.,
 Collymore, A., Cooke, P., Dearlano, R., Dewar, K., Diaz, J.S.,
 Dodge, S., Faro, S., Ferreira, P., Fitzhugh, W., Gage, D., Galagan, J.,
 Gardyna, S., Ginde, S., Goyette, M., Graham, L., Grand-Pierre, N.,
 Hages, B., Heaford, A., Horton, L., Hulme, W., Iliev, I., Johnson, R.,
 Jones, C., Karatas, A., Lakocque, K., Lamazares, R., Landers, T.,
 Lehoczký, J., Levine, R., Liu, G., Maclean, C., McDonald, P.,
 Marguies, N., Matthews, C., McCarthy, M., McEwan, P., McKernan, K.,
 McPheters, R., Meldrum, J., Menus, L., Mihova, T., Mlenga, V.,
 Murphy, T., Naylor, J., Nguyen, C., Norbu, C., Norman, C.H.,
 O'Connor, T., O'Donnell, F., O'Neill, D., Oliver, J., Peterson, K.,
 Phunkhang, P., Pierre, N., Pollara, V., Raymond, C., Retta, R.,
 Rieback, M., Riley, R., Rise, C., Rogov, P., Roman, J., Rosetti, M.,
 Roy, A., Santos, R., Schauer, S., Schuback, R., Seaman, S., Severy, P.,
 Sougnuez, C., Spencer, B., Stange-Thomann, N., Stojanovic, N.,
 Straus, N., Subramanian, A., Talamas, J., Teefaye, S., Theodore, J.,
 Travers, M., Travis, N., Trigilio, J., Vassiliev, H., Viel, R., Vo, A.,
 Wilson, B., Wu, X., Wyman, D., Ye, W.J., Young, G., Zainoun, J.,
 Zembek, L., Zimer, A. and Zody, M.
 Direct Submission
 Submitted (14-FEB-2001) Whitehead Institute/MIT Center for Genome
 Research, 320 Charles Street, Cambridge, MA 02141, USA
 3 (bases 1 to 172939)
 REFERENCES
 AUTHORS
 Birren, B., Nusbaum, C., Lander, E., Ali, A., Allen, N., Anderson, S.,
 Barna, N., Bastien, V., Bloom, T., Boguslavsky, L., Boukhalter, B.,
 Camarata, J., Chang, J., Chazaro, B., Choquet, Y., Collymore, A.,
 Cook, A., Cooke, P., Dearlano, K., Dewar, K., Diaz, J.S., Dodge, S.,
 Faro, S., Ferreira, P., FitzGerald, M., Gage, D., Galagan, J.,
 Gardyna, S., Gird, S., Graham, L., Grand-Pierre, N., Hages, B.,
 Horton, L., Hulme, W., Iliev, I., Johnson, R., Jones, C., Kamat, A.,
 Karatas, A., Kells, C., Landers, T., Levine, R., Lindblad-Toh, K.,
 Liu, G., Maclean, C., McDonald, P., Major, J., Matthews, C.,
 McCarthy, M., Meldrum, J., Menus, L., Mihova, T., Mlenga, V.,
 Murphy, T., Naylor, J., Nguyen, C., Nicol, R., Norbu, C., Norman, C.H.,
 O'Connor, T., O'Donnell, F., O'Neill, D., Oliver, J., Peterson, K.,
 Phunkhang, P., Pierre, N., Raymond, C., Retta, R., Rise, C., Rogov, P.,
 Roman, J., Roy, A., Schauer, S., Schuback, R., Seaman, S., Severy, P.,
 Smith, C., Spencer, B., Stange-Thomann, N., Stojanovic, N., Talamas, J.,
 Teefaye, S., Theodore, J., Topham, K., Travers, M., Vassiliev, H.,
 Viel, R., Vo, A., Wilson, B., Wu, X., Wyman, D., Young, G., Zainoun, J.,
 Zembek, L., Zimer, A. and Zody, M.
 Direct Submission
 Submitted (05-SEP-2002) Whitehead Institute/MIT Center for Genome
 Research, 320 Charles Street, Cambridge, MA 02141, USA
 4 (bases 1 to 172939)
 REFERENCES
 AUTHORS
 Birren, B., Nusbaum, C., Lander, E., Ali, A., Allen, N., Anderson, S.,
 Barna, N., Bastien, V., Bloom, T., Boguslavsky, L., Boukhalter, B.,
 Camarata, J., Chang, J., Chazaro, B., Choquet, Y., Collymore, A.,
 Cook, A., Cooke, P., Dearlano, K., Dewar, K., Diaz, J.S., Dodge, S.,
 Faro, S., Ferreira, P., FitzGerald, M., Gage, D., Galagan, J.,
 Gardyna, S., Gird, S., Graham, L., Grand-Pierre, N., Hages, B.,

Horton, L., Hulme, W., Iliev, I., Johnson, R., Jones, C., Kamat, A., Karatas, A., Kells, C., Landers, T., Levine, R., Lindblad-Toh, K., Liu, G., Maclean, C., Macdonald, P., Major, J., Matthews, C., McCarthy, M., Meldrum, J., Menees, L., Mihova, T., Mlenga, V., Murphy, T., Naylor, J., Nguyen, C., Nicol, R., Norbu, C., Norman, C.H., O'Connor, T., O'Donnell, P., O'Neill, D., Oliver, J., Peterson, K., Phunkhang, P., Pierre, N., Raymond, C., Retta, R., Rise, C., Rogov, P., Roman, J., Roy, A., Schauer, S., Schupbach, R., Seaman, S., Severy, P., Smith, C., Spencer, B., Strange-Thomann, N., Stojanovic, N., Talamas, J., Teefaye, S., Theodore, J., Topham, K., Travers, M., Vassiliev, H., Viel, R., Vo, A., Wilson, B., Wu, X., Wyman, D., Young, G., Zainoun, J., Zembek, L., Zimmer, A. and Zody, M.

Direct Submission
Submitted (01-OCT-2002) Whitehead Institute/MIT Center for Genome Research, 320 Charles Street, Cambridge, MA 02141, USA
On Oct 1, 2002 this sequence version replaced gi:22726159.
All repeats were identified using RepeatMasker:
Smit, A.F.A. & Green, P. (1996-1997)
http://ftp.genome.washington.edu/RM/RepeatMasker.html

Genome Center
Center: Whitehead Institute/ MIT Center for Genome Research
Center code: WIBR
Web site: http://www-seq.wi.mit.edu
Contact: sequence_submissions@genome.wi.mit.edu
----- Project Information
Center project name: 113_D_6
Center clone name: 113_D_6

----- Location/Qualifiers
1. 172939
/organism="Homo sapiens"
/mol_type="Genomic DNA"
/db_xref="taxon:9606"
/chromosome="11"
/map="11"
/clone="RP11-113D6"
/clone_11b="RP11-11 Human Male BAC"
repeat_region
complement(1..77)
/rpt_family="AluS"
complement(154..680)
/rpt_family="Charlie1"
complement(678..810)
/rpt_family="Charlie1"
complement(879..1057)
/rpt_family="L1M4"
complement(1395..1513)
repeat_region
/rpt_family="FLAM_A"
1870..1876
/note="<30 qual SNGL region"
1902..1910
/note="<30 qual SNGL region"
2672..2788
/rpt_family="SS"
complement(4009..4170)
/rpt_family="MIR"
complement(4210..4355)
/rpt_family="MERSB"
4356..4647
/rpt_family="AluSx"
complement(6554..6656)
/rpt_family="LTR41"
6916..6985
/rpt_family="MLT1C"
complement(7175..7359)
/rpt_family="LTR41"
complement(8308..8611)
repeat_region
/rpt_family="Aluub"
8622..9083
/rpt_family="MLT2B1"
complement(9265..10668)
repeat_region
/rpt_family="L1PB3"
10699..110609
/rpt_family="MER34"

repeat_region	10782..11264
repeat_region	/rpt_family="LTR48B"
repeat_region	complement(11344..11683)
repeat_region	/rpt_family="L2"
repeat_region	11839..12026
repeat_region	/rpt_family="L1MC4a"
repeat_region	12188..12660
repeat_region	/rpt_family="L1MA2"
repeat_region	12879..13322
repeat_region	/rpt_family="L1MC4a"
repeat_region	13223..13522
repeat_region	/rpt_family="AluSx"
repeat_region	13523..13817
repeat_region	/rpt_family="L1MC4a"
repeat_region	14819..15108
repeat_region	/rpt_family="AluJo"
repeat_region	15114..15175
repeat_region	/rpt_family="(GAA)n"
repeat_region	16150..16336
repeat_region	/rpt_family="MERSA"
repeat_region	complement(16587..16598)
repeat_region	/note="<30 qual SNGL region"
repeat_region	complement(16607..16612)
repeat_region	/note="<30 qual SNGL region"
repeat_region	complement(16746..17360)
repeat_region	/rpt_family="L1MB3"
repeat_region	17361..17664
repeat_region	/rpt_family="AluJo"
repeat_region	complement(17665..17774)
repeat_region	/rpt_family="L1MB3"
repeat_region	complement(17827..17929)
repeat_region	/rpt_family="MIR"
repeat_region	18993..19298
repeat_region	/rpt_family="AluJo"
repeat_region	19759..19821
repeat_region	/rpt_family="L2"
repeat_region	19933..19941
repeat_region	/note="<30 qual SNGL region"
repeat_region	20320..20609
repeat_region	/rpt_family="AluJo"
repeat_region	complement(20684..20994)
repeat_region	/rpt_family="AluSx"
repeat_region	20713..20721
repeat_region	/note="<30 qual SNGL region"
repeat_region	20879..20883
repeat_region	/note="<30 qual SNGL region"
repeat_region	complement(21118..21396)
repeat_region	/rpt_family="AluSx"
repeat_region	21141..21147
repeat_region	/note="<30 qual SNGL region"
repeat_region	21144
repeat_region	/note="probably A; possibly G"
repeat_region	21146
repeat_region	/note="probably A; possibly G"
repeat_region	21153..21158
repeat_region	/note="<30 qual SNGL region"
repeat_region	21154
repeat_region	/note="probably T; possibly C"
repeat_region	21160..21167
repeat_region	/note="<30 qual SNGL region"

Query Match 69.9%; Score 1830; DB 9; Length 172939;
Best Local Similarity 99.7%; Pred. No. 0;
Matches 2130; Conservative 0; Mismatches 6; Indels 0; Gaps 0;

QY	466	GGGTATCATGACTGGGTTTCTGAGCATGATTCATCAACCATCCAGTCTGGGTACAGAAC	525
DB	113998	GGGTATCATGACTGGGTTTCTGAGCATGATTCATCAACCATCCAGTCTGGGTACAGAAC	113939
QY	526	TGACACCATCAACGAGCGTGGAGAGACTCTGCTACAGACAGACCTGAGCTTACGG	585
DB	113938	TGACACCATCAACGAGCGTGGAGAGACTCTGCTACAGACAGACCTGAGCTTACGG	113879

QY 586 GAGTGAAGTCATCGTTCCCTTGTGCGCTGACAGAGAAACGCGGTTGTGCTGCTCC 645
DB 113878 GAGCTGAAGTCATCGTTCCCTTGTGCGCTGACAGAGAAACGCGGTTGTGCTGCTCC 113819
QY 646 TGGGCTGCGGCATGCGGAGAGAGGCTGTCTCCATCTACATCTCACTGCTGCGGCGG 705
DB 113818 TGGGCTGCGGCATGCGGAGAGAGGCTGTCTCCATCTACATCTCACTGCTGCGGCGG 113759
QY 706 ACTTCTCTCTCTTAAGGCGGCACATTAATGTTGCGCGTTACGCGCTCATCAATCTCCGC 765
DB 113758 ACTTCTCTCTCTTAAGGCGGCACATTAATGTTGCGCGTTACGCGCTCATCAATCTCCGC 113699
QY 766 ATCCCATCTTCCAAATCTCAGTCTGTGATGAACCTTCCCTAATTATAGGCGCTAAGCA 825
DB 113698 ATCCCATCTTCCAAATCTCAGTCTGTGATGAACCTTCCCTAATTATAGGCGCTAAGCA 113639
QY 826 TGTGTAGCGGCATCAGACCGAGGCGTGTCTGTCATCTGTGCGCATCTGTACCACT 885
DB 113638 TGTGTAGCGGCATCAGACCGAGGCGTGTCTGTCATCTGTGCGCATCTGTACCACT 113579
QY 886 GCGGCGCGCCCAATACCTGTCAATCGGTCAATGTTGTCGCTGTGCGCGCTGTCCCTGC 945
DB 113578 GCGGCGCGCCCAATACCTGTCAATCGGTCAATGTTGTCGCTGTGCGCGCTGTCCCTGC 113519
QY 946 TGGGAGTATCTGAGTGTGATGTTCTGTGACTTCTCTGTTAGTGTGTGATTTCTGTT 1005
DB 113518 TGGGAGTATCTGAGTGTGATGTTCTGTGACTTCTCTGTTAGTGTGTGATTTCTGTT 113459
QY 1006 GGTGTGAACGTCAGATTTCAATCAATGCGGTGTGTGTTTATATGTGTGTTCTCT 1065
DB 113458 GGTGTGAACGTCAGATTTCAATCAATGCGGTGTGTGTTTATATGTGTGTTCTCT 113399
QY 1066 GTGGGTTCAGGCTGTGTCTGTGTGAGATTCCTGTGTGATCCCGAAGATGCGGCTGA 1125
DB 113398 GTGGGTTCAGGCTGTGTCTGTGTGAGATTCCTGTGTGATCCCGAAGATGCGGCTGA 113339
QY 1126 CCAAGCTGTATCGTACCATCTCTCAAGTGTGTCTTCTCTCTGTGTGCGCTGCTC 1185
DB 113338 CCAAGCTGTATCGTACCATCTCTCAAGTGTGTCTTCTCTCTGTGTGCGCTGCTC 113279
QY 1186 TTGGCATTCAGTGGGCTCTGTTTCCAGATTCACCTGATTTGGAAGTCTTATTTGTTC 1245
DB 113278 TTGGCATTCAGTGGGCTCTGTTTCCAGATTCACCTGATTTGGAAGTCTTATTTGTTC 113219
QY 1246 ATGTGTATCTAGTTCATTTCTGTGCGCTCTTAACAGAGGCCAACCCCATCAATTT 1305
DB 113218 ATGTGTATCTAGTTCATTTCTGTGCGCTCTTAACAGAGGCCAACCCCATCAATTT 113159
QY 1306 ACTTCTTGTGAGGCTCTTATAGGAGGCTCAAAATAGGAGAACTGAGAGTGTCTCC 1365
DB 113158 ACTTCTTGTGAGGCTCTTATAGGAGGCTCAAAATAGGAGAACTGAGAGTGTCTCC 113099
QY 1366 AGAGGCTCTGACGACACGCTGTAGGTGATGAAGGTGAGAGGCTGTCTTCAAGAAA 1425
DB 113098 AGAGGCTCTGACGACACGCTGTAGGTGATGAAGGTGAGAGGCTGTCTTCAAGAAA 113039
QY 1426 CCTGTGAGCTGTGGGAGAGATTTGAGAGATGAGAGAGAACTCTGCGCTGTCAACA 1485
DB 113038 CCTGTGAGCTGTGGGAGAGATTTGAGAGATGAGAGAGAACTCTGCGCTGTCAACA 112979
QY 1486 GGAATTTGAGAGCAATGCTGCGCTGACCACTTGAACAATTAATGATTTTCTTACCT 1545
DB 112978 GGAATTTGAGAGCAATGCTGCGCTGACCACTTGAACAATTAATGATTTTCTTACCT 112919
QY 1546 TGTGCTCAGAAATGTCTCAGTGTCTCTCAAGTCTTGAATGATGTTATCTAACT 1605
DB 112918 TGTGCTCAGAAATGTCTCAGTGTCTCTCTCAAGTCTTGAATGATGTTATCTAACT 112859
QY 1606 GAGATTTGAGTTCATCCCATGAGAAAGCAATTAATGATGTTTGTGATTTCTC 1665
DB 112858 GAGATTTGAGTTCATCCCATGAGAAAGCAATTAATGATGTTTGTGATTTCTC 112799
QY 1666 CTGATATTAACAATATATTTCTGTATCTTGAACATGATCTTCTACTGAACT 1725

DB 112798 CTGATATTAACAATATATTTCTGTATCTTGAACATGATCTTCTACTGAACT 112739
QY 1726 TTTTTCGACATTTTCAATGTAATAAAGAGTGTGTGCAACAACCTTAAACCTTCT 1785
DB 112738 TTTTTCGACATTTTCAATGTAATAAAGAGTGTGTGCAACAACCTTAAACCTTCT 112679
QY 1786 TATATCTGTTCTTACCTGATGATGATCAAAAAGAGATTCCTTATTAATCTGTCAACT 1845
DB 112678 TATATCTGTTCTTACCTGATGATGATCAAAAAGAGATTCCTTATTAATCTGTCAACT 112619
QY 1846 ATGTTCCTGAAATATATGTTCCCTTATATGACTGAGGCAATTAATGAGTGTGAAGCT 1905
DB 112618 ATGTTCCTGAAATATATGTTCCCTTATATGACTGAGGCAATTAATGAGTGTGAAGCT 112559
QY 1906 CAATTCCTTATATGATGATCTGTACCTTAATTCATGAAATTCAGATTAAGC 1965
DB 112558 CAATTCCTTATATGATGATCTGTACCTTAATTCATGAAATTCAGATTAAGC 112499
QY 1966 AAATATAGACTTGAAGAGATTCCTCTCATTAATAACAGCTTGAATGTTT 2025
DB 112498 AAATATAGACTTGAAGAGATTCCTCTCATTAATAACAGCTTGAATGTTT 112439
QY 2026 ATGATAGCCCTCTCTGTCAATTTGTCACAGATGTGACATGTTGCTGTGTTCTA 2085
DB 112438 ATGATAGCCCTCTCTGTCAATTTGTCACAGATGTGACATGTTGCTGTGTTCTA 112379
QY 2086 GTAAAGCATGTGTGCGCTTCCCTGAGAACTGTATGTTTATTAATGCTCTCT 2145
DB 112378 GTAAAGCATGTGTGCGCTTCCCTGAGAACTGTATGTTTATTAATGCTCTCT 112319
QY 2146 GGAATATGAATAGTGAAGAGGCTTAATATATGTCACAGTTCATTTTGGCCATTG 2205
DB 112318 GGAATATGAATAGTGAAGAGGCTTAATATATGTCACAGTTCATTTTGGCCATTG 112259
QY 2206 GAAACTCAATATGATTTTAAAGTGAATTAATCTTGAACCAATTAATTAATCACTTA 2265
DB 112258 GAAACTCAATATGATTTTAAAGTGAATTAATCTTGAACCAATTAATTAATCACTTA 112199
QY 2266 CAGATCTTCAATGTAAGAGAAATCTTCATCTCCAGTTTGTATTAATGTTCTG 2335
DB 112198 CAGATCTTCAATGTAAGAGAAATCTTCATCTCCAGTTTGTATTAATGTTCTG 112139
QY 2326 ATGTATCTTCAATGTTTATGCTGTTTACATGAGAGCAAACTGAAACATCTG 2385
DB 112138 ATGTATCTTCAATGTTTATGCTGTTTACATGAGAGCAAACTGAAACATCTG 112079
QY 2386 ACCTTTCATGACATCTCAATTAATGATCTGATTAATACTTACAGTTGTACAGAT 2445
DB 112078 ACCTTTCATGACATCTCAATTAATGATCTGATTAATACTTACAGTTGTACAGAT 112019
QY 2446 TGTGATTCATGCTGTGACATCAATGAACCTGTGAATTTGTGCTAAGAGAAATAGCCAG 2505
DB 112018 TGTGATTCATGCTGTGACATCAATGAACCTGTGAATTTGTGCTAAGAGAAATAGCCAG 111959
QY 2506 ACGCAAAACATATGTGAATTCATTAATGATGATCAATTAATGAAATCTTGA 2565
DB 111958 ACGCAAAACATATGTGAATTCATTAATGATGATCAATTAATGAAATCTTGA 111899
QY 2566 CACGAAATTAATTAAGAGATCTGTGCTGGA 2601
DB 111898 CACGAAATTAATTAAGAGATCTGTGCTGGA 111863

RESULT 4
AC020568 192618 bp DNA linear HTG 17-AUG-2000
LOCUS AC020568
DEFINITION Homo sapiens chromosome 20 clone Rpl1-206cl, WORKING DRAFT
SEQUENCE 24 unordered pieces.
ACCESSION AC020568
VERSION AC020568.4 GI:9280789
KEYWORDS HTG; HTGS PHASBI; HTGS DRAFT.
SOURCE Homo sapiens (human)

ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Homidae; Homo.
REFERENCE 1 (bases 1 to 192618)
AUTHORS Waterston, R.H.
TITLE The sequence of Homo sapiens clone
JOURNAL Unpublished
2 (bases 1 to 192618)
REFERENCE Waterston, R.H.
TITLE Direct Submission
JOURNAL Submitted (04-JUN-2000) Genome Sequencing Center, Washington
University School of Medicine, 4444 Forest Park Parkway, St. Louis,
MO 63108, USA
On Jul 19, 2000 this sequence version replaced gi:9211362.
COMMENT ----- Genome Center -----
Center: Washington University Genome Sequencing Center
Web site: <http://genome.wustl.edu/gsc/index.shtml>
Project Information -----
Center project name: H NH0206C01
Summary Statistics -----
Sequencing vector: M13, 93%
Chemistry: Dye-terminator Big Dye, 7% of reads
Chemistry: Dye-terminator Big Dye, 7% of reads
Assembly program: Phrap; version 0.990319
Consensus quality: 179082 bases at least Q40
Consensus quality: 183120 bases at least Q40
Consensus quality: 185448 bases at least Q20
Insert size: 205000; agarose-fp
Insert size: 190318; sum-of-contigs
Quality coverage: 3.99 in Q20 bases; sum-of-contigs
Quality coverage: 4.34 in Q20 bases; sum-of-contigs

* NOTE: This is a 'working draft' sequence. It currently
* consists of 24 contigs. The true order of the pieces
* is not known and their order in this sequence record is
* arbitrary. Gaps between the contigs are represented as
* runs of N, but the exact sizes of the gaps are unknown.
* This record will be updated with the finished sequence
* as soon as it is available and the accession number will
* be preserved.
1 1279: contig of 1279 bp in length
* 1280 1379: gap of unknown length
* 1380 3544: contig of 2165 bp in length
* 3545 3644: gap of unknown length
* 3645 5441: contig of 1797 bp in length
* 5442 5541: gap of unknown length
* 5542 8421: contig of 2880 bp in length
* 8422 8521: gap of unknown length
* 8522 10096: contig of 1575 bp in length
* 10097 10196: gap of unknown length
* 10197 11938: contig of 1742 bp in length
* 11939 12038: gap of unknown length
* 12039 14708: contig of 2670 bp in length
* 14709 14808: gap of unknown length
* 14809 17876: contig of 3068 bp in length
* 17877 17976: gap of unknown length
* 17977 21534: contig of 3558 bp in length
* 21535 21634: gap of unknown length
* 21635 24615: contig of 2981 bp in length
* 24616 24715: gap of unknown length
* 24716 28301: contig of 3586 bp in length
* 28302 28401: gap of unknown length
* 28402 35969: contig of 7568 bp in length
* 35970 36069: gap of unknown length
* 36070 42473: contig of 6404 bp in length
* 42474 42573: gap of unknown length
* 42574 49752: contig of 7179 bp in length
* 49753 49852: gap of unknown length
* 49853 57604: contig of 7752 bp in length
* 57605 57704: gap of unknown length

FEATURES
source
1..192618
/organism="Homo sapiens"
/mol_type="genomic DNA"
/db_xref="taxon:9606"
/chromosome="20"
/clone="RP11-206C1"
1..1279
/note="assembly_name:Contig19"
1380..3544
/note="assembly_name:Contig21"
3645..5441
/note="assembly_name:Contig22"
5542..8421
/note="assembly_name:Contig23"
8522..10096
/note="assembly_name:Contig24"
10197..11938
/note="assembly_name:Contig25"
12039..14708
/note="assembly_name:Contig26"
14809..17876
/note="assembly_name:Contig27"
17977..21534
/note="assembly_name:Contig28"
21635..24615
/note="assembly_name:Contig29"
24716..28301
/note="assembly_name:Contig30"
28402..35969
/note="assembly_name:Contig31"
vector_end:SP6
vector_start:left
36070..42473
/note="assembly_name:Contig32"
42574..49752
/note="assembly_name:Contig33"
49853..57604
/note="assembly_name:Contig34"
57705..67088
/note="assembly_name:Contig35"
67189..77656
/note="assembly_name:Contig36"
77757..92143
/note="assembly_name:Contig37"
92240..109240
/note="assembly_name:Contig38"
109341..124285
/note="assembly_name:Contig39"
124386..138608
/note="assembly_name:Contig40"
138709..153882
/note="assembly_name:Contig41"
153983..169994
/note="assembly_name:Contig42"
170095..192618
Location/Qualifiers
57705 67088: contig of 9384 bp in length
* 67089 67188: gap of unknown length
* 67189 77656: contig of 10468 bp in length
* 77657 77757: gap of unknown length
* 77757 92143: contig of 14387 bp in length
* 92144 92243: gap of unknown length
* 92244 109240: contig of 16997 bp in length
* 109241 109341: gap of unknown length
* 109341 124285: contig of 14945 bp in length
* 124286 124385: gap of unknown length
* 124386 138608: contig of 14223 bp in length
* 138609 138708: gap of unknown length
* 138709 153882: contig of 15174 bp in length
* 153883 153982: gap of unknown length
* 153983 169994: contig of 16012 bp in length
* 169995 170094: gap of unknown length
* 170095 192618: contig of 22524 bp in length.

ORIGIN /note="assembly_name:Contig43"

Query Match 42.8%; Score 1120; DB 2; Length 192618;

Best Local Similarity 99.6%; Pred. No. 0;

Matches 1370; Conservative 0; Mismatches 5; Indels 0; Gaps 0;

```
QY 1227 TGGAAAGCTTATTTGTGATGATCTGATGTTCCATTTTCGTCGCGCTTAAGC 1286
DB 157148 TGGAAAGCTTATTTGTGATGATCTGATGTTCCATTTTCGTCGCGCTTAAGC 157207
QY 1287 AGTGCCAAACCCCAATCATTTACTTCTTCGTGCGCTCTTTAGGCAAGCTCAAAATAGCAG 1346
DB 157208 AGTGCCAAACCCCAATCATTTACTTCTTCGTGCGCTCTTTAGGCAAGCTCAAAATAGCAG 157267
QY 1347 AACCTGAAGCTGTTCTTCAGAGGCTCTGAGAGCAAGCTGCTGCAAGCTGATGAAGTGA 1406
DB 157268 AACCTGAAGCTGTTCTTCAGAGGCTCTGAGAGCAAGCTGCTGCAAGCTGATGAAGTGA 157327
QY 1407 GGGTGGCTTCTCAGAGAAACCTGAGAGCTGCGGAGAGCAATTGAGAGAGTGAAGAA 1466
DB 157328 GGGTGGCTTCTCAGAGAAACCTGAGAGCTGCGGAGAGCAATTGAGAGAGTGAAGAA 157387
QY 1467 ACCCTGCGCTCTCAGAGAGCAAGCTTGAAGCAATGCTGCTGCAAGCTTGAAGTGA 1526
DB 157388 ACCCTGCGCTCTCAGAGAGCAAGCTTGAAGCAATGCTGCTGCAAGCTTGAAGTGA 157447
QY 1527 TATGCAATTTTCTTACGCTTCTGCTGCAAGATGCTCAGTGGTCCCTCAAGTCTTGA 1586
DB 157448 TATGCAATTTTCTTACGCTTCTGCTGCAAGATGCTCAGTGGTCCCTCAAGTCTTGA 157507
QY 1587 ATGAGATTTATCTAACCCTGAGCTGAGCTTTCACCCAGTGAAGCAATTAAGCTGACA 1646
DB 157508 ATGAGATTTATCTAACCCTGAGCTGAGCTTTCACCCAGTGAAGCAATTAAGCTGACA 157567
QY 1647 GTACAAATGTTGATCTCCTGATATTAACAATACATTTTCCCTGATCTTGAAGTGA 1706
DB 157568 GTACAAATGTTGATCTCCTGATATTAACAATACATTTTCCCTGATCTTGAAGTGA 157627
QY 1707 ATCTTCTCACTGAGACATTTTCTGCACTTTTCATTTGATTAATAAGAGTGTCTGCA 1766
DB 157628 ATCTTCTCACTGAGACATTTTCTGCACTTTTCATTTGATTAATAAGAGTGTCTGCA 157687
QY 1767 CAACCCCTTAAACCTTCTTATATCTGTTTCTTACCTGATAGTGAAGAGTTC 1826
DB 157688 CAACCCCTTAAACCTTCTTATATCTGTTTCTTACCTGATAGTGAAGAGTTC 157747
QY 1827 CTATTTATCTGTCAGACTATGTTCCCTGAAATCATGTTCCCTTTATGACTGAGGC 1886
DB 157748 CTATTTATCTGTCAGACTATGTTCCCTGAAATCATGTTCCCTTTATGACTGAGGC 157807
QY 1887 ATTACTGAGCTGAGAGCTCAATTTCTTAATAGTGAAGTGTCTGACTTCAATTCAT 1946
DB 157808 ATTACTGAGCTGAGAGCTCAATTTCTTAATAGTGAAGTGTCTGACTTCAATTCAT 157867
QY 1947 GAATTTCTCAATTAATAAGCAAAATTAAGACTTGAAGAGATTCCTCCCTCAATAAAC 2006
DB 157868 GAATTTCTCAATTAATAAGCAAAATTAAGACTTGAAGAGATTCCTCCCTCAATAAAC 157927
QY 2007 AGCTTGAAGAAATGTTTATGAATAGCCCTCTCTCATTTTTCACAGACTGAGC 2066
DB 157928 AGCTTGAAGAAATGTTTATGAATAGCCCTCTCTCATTTTTCACAGACTGAGC 157987
QY 2067 ATGTGGCTTGTGTTCTTAAGTAAGCAATGTGGCCCTTCCCTTGAAGACTGTAAG 2126
DB 157988 ATGTGGCTTGTGTTCTTAAGTAAGCAATGTGGCCCTTCCCTTGAAGACTGTAAG 158047
QY 2127 TTCTTATTTAGCTTCTTCTGAGCAATGAAGTGAAGAGCTTAAATATGTCACACC 2186
DB 158048 TTCTTATTTAGCTTCTTCTGAGCAATGAAGTGAAGAGCTTAAATATGTCACACC 158107
QY 2187 AGTTTCATTTGGCATTTGAGAACTCAATATGATTTTAAAGTGAATTAATCTTGAAA 2246
DB 158107 AGTTTCATTTGGCATTTGAGAACTCAATATGATTTTAAAGTGAATTAATCTTGAAA 158167
```

```
DB 158108 AGTTTCATTTGGCATTTGAGAACTCAATATGATTTTAAAGTGAATTAATCTTGAAA 158167
QY 2247 ACCATTTATTTACACTTACAGATTCCTTTCAGTTGAGAGAAATCTTCACTTCCAGG 2306
DB 158168 ACCATTTATTTACACTTACAGATTCCTTTCAGTTGAGAGAAATCTTCACTTCCAGG 158227
QY 2307 TTTTGTATTAATTTGTTCTGATTTGAATCTTTCAGTTAGTTTATGCTGTTTACATGAAA 2366
DB 158228 TTTTGTATTAATTTGTTCTGATTTGAATCTTTCAGTTAGTTTATGCTGTTTACATGAAA 158287
QY 2367 GCAAACTGAAAATCTGACTTTCATGACATCTCAATTAATGATCTGATTAATA 2426
DB 158288 GCAAACTGAAAATCTGACTTTCATGACATCTCAATTAATGATCTGATTAATA 158347
QY 2427 CTTCAGTTGTTGATGATTTCTGATTAATCTTTCAGTTAGTTTATGCTGTTTACATGAAA 2486
DB 158348 CTTCAGTTGTTGATGATTTCTGATTAATCTTTCAGTTAGTTTATGCTGTTTACATGAAA 158407
QY 2487 GCTAAGGAAATTAAGCAAGAGCCAAACATATTTGTAAGTTCAATTTCTAAGATATCC 2546
DB 158408 GCTAAGGAAATTAAGCAAGAGCCAAACATATTTGTAAGTTCAATTTCTAAGATATCC 158467
QY 2547 AAATTAAGAAATTTCTTGAACAAGAAATTAATTAAGAGATCTGCTGGA 2601
DB 158468 AAATTAAGAAATTTCTTGAACAAGAAATTAATTAAGAGATCTGCTGGA 158522

RESULT 5
AX429467 8622 bp DNA linear PAT 21-JUN-2002
LOCUS AX429467
DEFINITION Sequence 3 from Patent WO0234914.
ACCESSION AX429467
VERSION AX429467.1 GI:21540764
KEYWORDS
SOURCE Homo sapiens (human)
ORGANISM Homo sapiens
            Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.
REFERENCE 1
AUTHORS Zhao,Q., Beasley,E.M. and Wei,M.H.
TITL Isolating human g-protein coupled receptors, nucleic acid molecules
        encoding human gpcr proteins, and uses thereof
JOURNAL Patent: WO 0234914-A 3 02-MAY-2002;
        PE CORP NY (US)
FEATURES
        Location/Qualifiers
        source
            1..8622
            /organism="Homo sapiens"
            /mol_type="unassigned DNA"
            /db_xref="taxon:9606"
ORIGIN
Query Match 42.1%; Score 1101; DB 6; Length 8622;
Best Local Similarity 100.0%; Pred. No. 0;
Matches 1101; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 466 GGGTCATCAGACCTGGGTTCTGAGCATGATTAACCATCCAGTGTGGGTACAGAAC 525
DB 7502 GGGTCATCAGACCTGGGTTCTGAGCATGATTAACCATCCAGTGTGGGTACAGAAC 7551
QY 526 TGAACCAATCAACGAGCTGAGAGACTCTTGTCTCAAGAGACCCCTGAGCTTCAAG 585
DB 7562 TGAACCAATCAACGAGCTGAGAGACTCTTGTCTCAAGAGACCCCTGAGCTTCAAG 7621
QY 586 GGGTCATCAGACCTGGGTTCTGAGCATGATTAACCATCCAGTGTGGGTACAGAAC 645
DB 7622 GGGTCATCAGACCTGGGTTCTGAGCATGATTAACCATCCAGTGTGGGTACAGAAC 7681
QY 646 TGGCTGCGGCAATGCGAGAGAGCTGTCATCATTAATCAATCTCAACCTGCTGCGGCG 705
DB 7682 TGGCTGCGGCAATGCGAGAGAGCTGTCATCATTAATCAATCTCAACCTGCTGCGGCG 7741
QY 706 ACTTCTCTTCTTACGCGGCAATTAATGATTTGCGCGTTACGCTCATCAATATCCGCC 765
DB 7741 ACTTCTCTTCTTACGCGGCAATTAATGATTTGCGCGTTACGCTCATCAATATCCGCC 158167
```


ORIGIN

Query Match 42.0%; Score 1100; DB 9; Length 1697;
 Best Local Similarity 99.7%; Pred. No. 0;
 Matches 1300; Conservative 0; Mismatches 4; Indels 0; Gaps 0;

/translation="MDSTIPVLGTELPINGREBPCYKOTLSPTGLTCTIVSLVALTG
 NAVWLMLGCRMRNNAVSIYILMLVADPFLSGHILRSPLRLINHPISKLSPFM
 TPEYFGLSMLSAISTERCISILMPYHCRPRYSVVCVLMALSLRSLIEMWF
 CDEPLFGSANSVCESTDPIFIAMLVFCVLCSSVLVRLIIGSRKMPILTYVYI
 LTLVLVFLCGLPFGIOMALPSRIHIDMKVLPCHVHVSIFLSALNSANPIIYFVVG
 SFRQRNRQMLKVLQRLADTPEVDEGGWLPQETLELSGRLEQ"

466 GGGTCATCAGACTGGGGTTTCTAGCATGATTCACCATCCAGTCTGGGTACAGAAC 525
 336 GGGTCATCAGACTGGGGTTTCTAGCATGATTCACCATCCAGTCTGGGTACAGAAC 395
 526 TGACCAATCAACGAGCGTAGAGAGACTCTTGTCTACAGAGACCTTGACTTCACGG 585
 396 TGACCAATCAACGAGCGTAGAGAGACTCTTGTCTACAGAGACCTTGACTTCACGG 455
 586 GGGTCAGTGCATCGTTCCCTGTGCGCTGACAGAGAAAGCGGTGTGTGCGCTCC 645
 456 GGGTCAGTGCATCGTTCCCTGTGCGCTGACAGAGAAAGCGGTGTGTGCGCTCC 515
 646 TGGGCTGCGCATGCGCAGAGAGCGTGTCTCATCTACATCTTCACCTGGTCCGCG 705
 516 TGGGCTGCGCATGCGCAGAGAGCGTGTCTCATCTACATCTTCACCTGGTCCGCG 575
 706 ACTTCTCTTCTTACGCGCAATTAATGTTGCGCGTTACGCTCATCAATATCCGCC 765
 576 ACTTCTCTTCTTACGCGCAATTAATGTTGCGCGTTACGCTCATCAATATCCGCC 635
 766 ATCCCATCTCCAAATCTCAAGTCTGTGAGAGACCTTCCCTTAATAGGCGCTACGA 825
 636 ATCCCATCTCCAAATCTCAAGTCTGTGAGAGACCTTCCCTTAATAGGCGCTACGA 695
 826 TGGTGAAGCGCATCAGACCGAGCGTGTCTCATCTGTGAGCGCATCTGTAGCACT 885
 696 TGGTGAAGCGCATCAGACCGAGCGTGTCTCATCTGTGAGCGCATCTGTAGCACT 755
 886 GCGCGCGCCCAAGATCCTGTCTATGCTGTCTGTCTGTGCGCGCTGTCTCTGC 945
 756 GCGCGCGCCCAAGATCCTGTCTATGCTGTCTGTCTGTGCGCGCTGTCTCTGC 815
 946 TGGGAGTATCTCGAGTGTGATGTTCTGTGATCTTCTGTGAGTGTGATCTCTCT 1005
 816 TGGGAGTATCTCGAGTGTGATGTTCTGTGATCTTCTGTGAGTGTGATCTCTCT 875
 1006 GGTGTGAACGTCAGATTTTCAATCAATCGCGTGTGATTTTATGTTGTTCTCT 1065
 876 GGTGTGAACGTCAGATTTTCAATCAATCGCGTGTGATTTTATGTTGTTCTCT 935
 1066 GTGGGTTCAGCTGTGTCTGTGTGAGATTTCTGTGATTCGCGAAGATGCGCTGA 1125
 936 GTGGGTTCAGCTGTGTCTGTGTGAGATTTCTGTGATTCGCGAAGATGCGCTGA 995
 1126 CGAGGCTGTACGTACCACTCTCTCAAGTGTGTCTCTCTCTGTGCGCTGTGCCT 1185
 996 CGAGGCTGTACGTACCACTCTCTCAAGTGTGTCTCTCTCTGTGCGCTGTGCCT 1055
 1186 TTGGGATTCAGTGGGCGCTGTTCAGAGATTCACCTGATTTGAAAGTCTTATTTTGT 1245
 1056 TTGGGATTCAGTGGGCGCTGTTCAGAGATTCACCTGATTTGAAAGTCTTATTTTGT 1115
 1246 ATGTGATCTAGTTCATTTTCTGTCTGTCTTAAACAGAGTGCACACCATCATTT 1305
 1116 ATGTGATCTAGTTCATTTTCTGTCTGTCTTAAACAGAGTGCACACCATCATTT 1175
 1306 ACTTCTCTGTGCGCTCTTAAAGGAGGTCAAAATAGGAGAACCTGAAGTGTCTCC 1365
 1176 ACTTCTCTGTGCGCTCTTAAAGGAGGTCAAAATAGGAGAACCTGAAGTGTCTCC 1235
 1366 AGAGGCGCTGTGAGAGACGCGCTGAGTGTGAAGGTGAGGCTGTCTCTCAGAGAA 1425

|||||
 Db 1236 AGAGGCGCTGTGAGAGACGCGCTGAGGTGATGAGGTGAGAGGCGACTTCTCAGAGAA 1295
 1426 CCTGTGAGCGTGTGGGAGAGAGATTGGAGAGAGAGAGAACCTGTGCCCTGTACAGA 1485
 1296 CCTGTGAGCGTGTGGGAGAGAGATTGGAGAGAGAGAGAACCTGTGCCCTGTACAGA 1355
 1486 GCACTTTGAGAGAGAGATGCTGCGTGCACCTTGACATTAATATGATTTTCTAGCCT 1545
 1356 GCACTTTGAGAGAGAGATGCTGCGTGCACCTTGACATTAATATGATTTTCTAGCCT 1415
 1546 TGTGCTCAGAAATGTCTCAGTGTCTCTCAAGTCTTGCATATGATTTTCTAGCCT 1605
 1416 TGTGCTCAGAAATGTCTCAGTGTCTCTCAAGTCTTGCATATGATTTTCTAGCCT 1475
 1606 GACAGTTCAGTTTTCACCCATGAGAAAGCATTTAGTCTGACAGTACAAATGTTGATCTC 1665
 1476 GACAGTTCAGTTTTCACCCATGAGAAAGCATTTAGTCTGACAGTACAAATGTTGATCTC 1535
 1666 CTGATATTAACAATATTTTCTGTTATCTTGTGACATGATCTTCTACTGACACT 1725
 1536 CTGATATTAACAATATTTTCTGTTATCTTGTGACATGATCTTCTACTGACACT 1595
 1726 TTTTCTGACATTTTCAATTTGATTAATAAGAGTGTGTGCAAA 1769
 1596 TTTTCTGACATTTTCAATTTGATTAATAAGAGTGTGTGCAAA 1639

RESULT 7
 AX647081 1369 bp DNA linear PAT 04-MAR-2003
 LOCUS Sequence 1273 from Patent EP1270724.
 DEFINITION AX647081
 ACCESSION AX647081
 VERSION AX647081.1 GI:28800068
 KEYWORDS
 SOURCE
 ORGANISM
 Homo sapiens (human)
 Homo sapiens
 Burkayota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.
 REFERENCE
 1 Suwa, M., Asai, K., Akiyama, Y. and Aburatani, H.
 TITLE Guanine triphosphate-binding protein coupled receptors
 JOURNAL Patent: EP 1270724-A 1273 02-JAN-2003;
 National Institute of Advanced Industrial Science and Technology
 (JP) ; Center for Advanced Science and Technology Incubation, Ltd.
 (JP)
 FEATURES
 source
 1. 1369
 /organism="Homo sapiens"
 /mol_type="unassigned DNA"
 /db_xref="taxon:9606"
 201..1169
 CDS
 /note="unnamed protein product"
 /codon_start=1
 /protein_id="CAD69666.1"
 /db_xref="GI:28800069"
 /translation="MDSTIPVLGTELPINGREBPCYKOTLSPTGLTCTIVSLVALTG
 NAVWLMLGCRMRNNAVSIYILMLVADPFLSGHILRSPLRLINHPISKLSPFM
 TPEYFGLSMLSAISTERCISILMPYHCRPRYSVVCVLMALSLRSLIEMWF
 CDEPLFGSANSVCESTDPIFIAMLVFCVLCSSVLVRLIIGSRKMPILTYVYI
 LTLVLVFLCGLPFGIOMALPSRIHIDMKVLPCHVHVSIFLSALNSANPIIYFVVG
 SFRQRNRQMLKVLQRLADTPEVDEGGWLPQETLELSGRLEQ"

ORIGIN

Query Match 41.7%; Score 1093; DB 6; Length 1369;
 Best Local Similarity 99.8%; Pred. No. 0;
 Matches 1193; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

466 GGGTCATCAGACTGGGGTTTCTAGCATGATTCACCATCCAGTCTGGGTACAGAAC 525
 175 GGGTCATCAGACTGGGGTTTCTAGCATGATTCACCATCCAGTCTGGGTACAGAAC 234
 526 TGACCAATCAACGAGCGTAGAGAGACTCTTGTCTACAGAGACCTTGAGCTTCACGG 585

Db	235	TGACACCAATCAAACGACCGTAGAGAACTCTTGCTACAAAGACACCTTAGCTTACACGG	294
Qy	586	GGCTGACGTGCATCGTTTTCCCTTGTCGCGCTGACAGAGAAACGGGGTGTGTCTCTGACCTCC	645
Db	295	GGCTGACGTGCATCGTTTTCCCTTGTCGCGCTGACAGAGAAACGGGGTGTGTCTCTGACCTCC	354
Qy	646	TGGGCTGCCGCATATGCCAGAGAAACGTGTCTCCATCTACATCTCTCAACCTGTGTGCCGGCG	705
Db	355	TGGGCTGCCGCATATGCCAGAGAAACGTGTCTCCATCTACATCTCTCAACCTGTGTGTGCCGGCG	414
Qy	706	ACTTCCTCTTCTCTTAGCGGCGACATATATATGTTCGGCGTTAAGCCTCATCATATATCGCGC	765
Db	415	ACTTCCTCTTCTCTTAGCGGCGACATATATATGTTCGGCGTTAAGCCTCATCATATATCGCGC	474
Qy	766	ATCCCATCTCCAAATATCTCAAGTCCGTGTATGACCTTTCCCTAATTATATAGCCTAAGCA	825
Db	475	ATCCCATCTCCAAATATCTCAAGTCCGTGTATGACCTTTCCCTAATTATATAGCCTAAGCA	534
Qy	826	TGCTGAGCGCCCATCAGACACGAGCGCTGCTGTCCATCTGTGTGCCCATCTGTGTACCACT	885
Db	535	TGCTGAGCGCCCATCAGACACGAGCGCTGCTGTCCATCTGTGTGCCCATCTGTGTACCACT	594
Qy	886	GGCGCGCGCCCGAATACCTGTGATATGGATATGTGTTCGTTCGTCTGTGGGCGTGTCCCTGC	945
Db	595	GGCGCGCGCCCGAATACCTGTGATATGGATATGTGTTCGTTCGTCTGTGGGCGTGTCCCTGC	654
Qy	946	TGCGAGATATCCTGTGAGTGAATGTCTGTGACCTTCCGTATATAGTGTGCTGATTTCTGTTT	1001
Db	655	TGCGAGATATCCTGTGAGTGAATGTCTGTGACCTTCCGTATATAGTGTGCTGATTTCTGTTT	714
Qy	1006	GGTGTGAAAAGTTCAGATTTTCATTTACATATGCGGTGCTGTTTTTATATGTGTGTCTCT	106
Db	715	GGTGTGAAAAGTTCAGATTTTCATTTACATATGCGGTGCTGTTTTTATATGTGTGTCTCT	774
Qy	1066	GTGGGTCACGCTGGTCTGCTGTGTAGAGATTTCTGTGTGATATCCGGAGAAATGCGGCTGA	112
Db	775	GTGGGTCACGCTGGTCTGCTGTGTAGAGATTTCTGTGTGATATCCGGAGAAATGCGGCTGA	834
Qy	1126	CCAGGCTGTACGTGACCAATCTCTCTCAACAGTGTGTCTTCTCTCTGTGTGCGTCCCT	118
Db	835	CCAGGCTGTACGTGACCAATCTCTCTCAACAGTGTGTCTTCTCTCTGTGTGCGTCCCT	894
Qy	1186	TTGGCATTCAGTGTGGGCGCTGTGTTTCCAGGATTCACCTGTATGGAAAAGTCTTATTTTGTTC	124
Db	895	TTGGCATTCAGTGTGGGCGCTGTGTTTCCAGGATTCACCTGTATGGAAAAGTCTTATTTTGTTC	954
Qy	1246	ATGTGCATTCATGTTTTCATTTTTCCTGTCCGCTCTTAAACAGCAGTGTCCAAACCCCATCATTT	130
Db	955	ATGTGCATTCATGTTTTCATTTTTCCTGTCCGCTCTTAAACAGCAGTGTCCAAACCCCATCATTT	101
Qy	1306	ACTTCTTTCGTGTGGGCTCTTTAAGGCAAGCTCAAAATATAGGCAAGCTGAAAGTGTCTCC	136
Db	1015	ACTTCTTTCGTGTGGGCTCTTTAAGGCAAGCTCAAAATATAGGCAAGCTGAAAGTGTCTCC	107
Qy	1366	AGAGGAGCTGTGCAGACACAGCCTGTAGAGTGAAGAGTGAAGGGGTGCTTCTCAAGAAA	142
Db	1075	AGAGGAGCTGTGCAGACACAGCCTGTAGAGTGAAGAGTGAAGGGGTGCTTCTCAAGAAA	113
Qy	1426	CCCTGTGAGCTGTTCGGAGAACGATTTGGAGCAGTGAAGAAAGAACTCTGCCCTGTCAAGCA	148
Db	1135	CCCTGTGAGCTGTTCGGAGAACGATTTGGAGCAGTGAAGAAAGAACTCTGCCCTGTCAAGCA	119
Qy	1486	GGACTTTTGAGAGCAATGTCCGCTGTGCAACCTTGAAGCAATTAATATGCAATTTTGTAGGCT	154
Db	1195	GGACTTTTGAGAGCAATGTCCGCTGTGCAACCTTGAAGCAATTAATATGCAATTTTGTAGGCT	125
Qy	1546	TCTGCCTCAGAAATGTCTCAGTGTGTCCCTCAAGGTCTTTCGAATAGATGTTTATCTAACT	160
Db	1255	TCTGCCTCAGAAATGTCTCAGTGTGTCCCTCAAGGTCTTTCGAATAGATGTTTATCTAACT	131
Qy	1606	CACAGTTGCAGTTTTTACCCATGGAAAAGCATTAAGTCTGACAGTACATGTTTTGGA	1660

DB	LOCUS	SEQUENCE	LENGTH	DNA	LINEAR	PAT
Db	1315	GACAGTTCGACGTTTCCACCACCGAAGAAGCATTAGCTGACAGTCAAGTGTGGGA	1369			
RESULT 8						
AX299158	Sequence 30 from Patent WO0183555.		1400 bp			
LOCUS	AX299158					
ACCESSION	AX299158					
VERSION	GI:17129065					
KEYWORDS						
SOURCE	Homo sapiens (human)					
ORGANISM	Homo sapiens					
REFERENCE	Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;					
AUTHORS	Mammalia; Eutheria; Primates; Catarrhini; Hominiidae; Homo.					
TITLE	Anderson, D.J., Dong, X., Zylka, M., Han, S.K. and Simon, M.					
JOURNAL	Pain signaling molecules					
FEATURES	Patent: WO 0183555-A 30 08-NOV-2001;					
SOURCE	CALIFORNIA INSTITUTE OF TECHNOLOGY (US)					
Location/Qualifiers						
1..1400						
/organism="Homo sapiens"						
/mol_type="unassigned DNA"						
/db_xref="taxon:9606"						
332..1300						
/note="unnamed protein product"						
/codon_start=1						
/protein_id="CAD12713.1"						
/db_xref="GI:17129066"						
/translation="MDSTIPVLGTELPIINGREETPCYKOTLSPFTGLCTIVALTGTA						
NAVVLWILGCMRNVAISYIINLVADPLFLSGHIICSLRLINIRHPISKLSPPVM						
TFEYFGLSMISAISTERCISILMPHYTHRRPYSYWCVLMALSLSRIEMNF						
CDLFPSADSWECSTDFITLAVLPCLVDCSSLVLLVLCGSRMPTLRIVTVII						
LTLVLPVLCGLPEIGIOVALFSRILHDMKVLFCHVHLVSIFLSLNSANDIYFEVG						
SFRORNRONIKLVIALQDTPDEVDEGGMLPETLELSGRLEQ"						
ORIGIN						
Query Match	41.6%; Score 1088; DB 6; Length 1400;					
Best Local Similarity	100.0%; Pred. No. 0;					
Matches 1088; Conservative 0; Mismatches 0; Indels 0; Gaps 0;						
Oy	473	CAGACTGGGGTTTCAGACATGATGAAACCATCCAGCTCTTGAGTACAGAACGACACC	532			
Db	313	CAGACTGGGGTTTCAGACATGATGAAACCATCCAGCTCTTGAGTACAGAACGACACC	372			
Oy	533	AATCAACGACGATGAGAGACTCTCTTGCTACAACACAGACCCCTGAGCTTCAAGGGAGTAC	592			
Db	373	AATCAACGACGATGAGAGACTCTCTTGCTACAACACAGACCCCTGAGCTTCAAGGGAGTAC	432			
Oy	593	GTGCATCGTTTTCCCTTGTGCGGCTGACAGAAAACGGGTTGTGCTCTGAGCTCTGGGCTG	652			
Db	433	GTGCATCGTTTTCCCTTGTGCGGCTGACAGAAAACGGGTTGTGCTCTGAGCTCTGGGCTG	492			
Oy	653	CCGAGTGGGAGAAAGCTGTCTCACTCACTCAATCTCAACCTGATCCGGGCGGACCTTCT	712			
Db	493	CCGAGTGGGAGAAAGCTGTCTCACTCACTCAATCTCAACCTGATCCGGGCGGACCTTCT	552			
Oy	713	CTTCCTTAGCGGACATTATATGTTGCGCGTTACGCCCTCATCAATAATCCGCATCCCAT	772			
Db	553	CTTCCTTAGCGGACATTATATGTTGCGCGTTACGCCCTCATCAATAATCCGCATCCCAT	612			
Oy	773	CTCCAAATCTCAGTCTCTGTGATGACCTTCCCTTACTTTATAGGCGCTAAGACGTGAG	832			
Db	613	CTCCAAATCTCAGTCTCTGTGATGACCTTCCCTTACTTTATAGGCGCTAAGACGTGAG	672			
Oy	833	CGGCATCAGACCGAGAGGCTGCTGTCATCTGTGGCCCATCTGATACCATCTGCGCGG	892			
Db	673	CGGCATCAGACCGAGAGGCTGCTGTCATCTGTGGCCCATCTGATACCATCTGCGCGG	732			
Oy	893	CCCCAGATACCTGATCGATCGATGTGATCTGCTGAGGCGGCTGCTGCTGCGGAG	952			
Db	733	CCCCAGATACCTGATCGATCGATGTGATCTGCTGAGGCGGCTGCTGCTGCGGAG	792			

QY 1553 CAGAAATG 1560
Db 1393 CAGAAATG 1400

RESULT 10
AY042215 1400 bp DNA linear PRI 14-SEP-2001
LOCUS Homo sapiens G protein-coupled receptor (MRGX3) gene, complete cds.
DEFINITION
ACCESSION AY042215
VERSION AY042215.1 GI:15546065
KEYWORDS
SOURCE Homo sapiens (human)
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Homidae; Homo.
REFERENCE
AUTHORS Dong, X., Han, S., Zylka, M.J., Simon, M.I. and Anderson, D.J.
TITLE A diverse family of G-proteins expressed in specific subsets of nociceptive sensory neurons
JOURNAL Cell 106 (5), 619-632 (2001)
MEDLINE 21435808
PUBMED 11551509
REFERENCE
AUTHORS Dong, X., Han, S., Zylka, M.J., Simon, M.I. and Anderson, D.J.
TITLE Direct Submission
JOURNAL Submitted (19-JUN-2001) Division of Biology, 216-76, California Institute of Technology, 1201 E. California Blvd., Pasadena, CA 91125, USA

FEATURES
Source
1. 1400
/organism="Homo sapiens"
/mol_type="genomic DNA"
/db_xref="taxon:9606"
<332..>1300
/gene="MRGX3"
<332..>1300
/gene="MRGX3"
/product="G protein-coupled receptor"
332..1300
/gene="MRGX3"
/note="Mab-related gene X3"
/codon_start=1
/product="G protein-coupled receptor"
/protein_id="AAK91806.1"
/db_xref="GI:15546065"
translation="MDSRIPIVLELPINGBETPCYKOTLSFGLTCIYSLVLTG
NAVYLMILGCMRNNAVITLNLVYADPLSGHITCSPLRINIRHPSKILSPLM
TPPYFGLSMLSAISTERCSTLMPFTHCRPRYSVMCYLMALSLSILEMFM
CDLFLSGADSVWCSTSDFTTAMLVFLCVGLSVLVLVLLGSRMPLRLVTVI
LTLVLPFLCGLPGIOWALPFRHLDKVLFCVHALVSLPLSALNSSANPLTFYFG
SFRORNRONIKVLVGRALQDTPVEDBGGMLPRTLLSGSRLEQ"

ORIGIN
Query Match 41.6%; Score 1088; DB 9; Length 1400;
Best Local Similarity 100.0%; Pred. No. 0;
Matches 1088; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 473 CAGACCTGGGCTTCTGACATGATTAACATCCAGCTGGGTACAGAACTGACACC 532
Db 313 CAGACCTGGGCTTCTGACATGATTAACATCCAGCTGGGTACAGAACTGACACC 372

QY 533 AATCAAGCAGAGCTGAGAGACTCTTGTCTACAGAGAGACCTGAGCTTCAACGGGGCTGAC 592
Db 373 AATCAAGCAGAGCTGAGAGACTCTTGTCTACAGAGAGACCTGAGCTTCAACGGGGCTGAC 432

QY 593 GTGCATGCTTCCCTTGTGCGCTGACAGAAACCGGGTGTGCTCTGCGCTCTTGGGCTG 652
Db 433 GTGCATGCTTCCCTTGTGCGCTGACAGAAACCGGGTGTGCTCTGCGCTCTTGGGCTG 492

QY 653 CCGCATGGGAGAAACGCTGTCTCACTACATCTCTCAACCGGTGCGGGCGAGCTTCTCT 712
Db 493 CCGCATGGGAGAAACGCTGTCTCACTACATCTCTCAACCGGTGCGGGCGAGCTTCTCT 552

QY 713 CTTCCTTAGCGGCCACATTAATATGTTGCGCGTTACGCTCTCATCAATATCCGCCATCCCAT 772
Db 553 CTTCCTTAGCGGCCACATTAATATGTTGCGCGTTACGCTCTCATCAATATCCGCCATCCCAT 612

QY 773 CTCGAAATCTCTGAGTCCCTGATGAGCTTCCCTACTTTATAGGCTTAAGCAATGCTGAG 832
Db 613 CTCGAAATCTCTGAGTCCCTGATGAGCTTCCCTACTTTATAGGCTTAAGCAATGCTGAG 672

QY 833 CGCCATCAGACACGAGCGCTGCTGCTCATCTGTCGAGCCCATCTGTATACACATGCGCGG 892
Db 673 CGCCATCAGACACGAGCGCTGCTGCTCATCTGTCGAGCCCATCTGTATACACATGCGCGG 722

QY 893 CCCCAGATACCTGTCAATGCTGATGATGATGCTCTGCTGAGCCCTGCTGCTGCGGAG 952
Db 723 CCCCAGATACCTGTCAATGCTGATGATGATGCTCTGCTGAGCCCTGCTGCTGCGGAG 792

QY 953 TATCCTGAGAGGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 1012
Db 793 TATCCTGAGAGGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 852

QY 1013 AACGTCAGATTTCAATTAATCAATCGAGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 1072
Db 853 AACGTCAGATTTCAATTAATCAATCGAGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 912

QY 1073 CAGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 1132
Db 913 CAGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 972

QY 1133 GTAGTGACATGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 1192
Db 973 GTAGTGACATGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 1032

QY 1193 TCAGTGAGCCCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 1252
Db 1033 TCAGTGAGCCCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 1092

QY 1253 TCTAGTTTCCATTTTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 1312
Db 1093 TCTAGTTTCCATTTTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 1152

QY 1313 CGTGCGGCTCTTTAGGCGAGCTGCAAAATAGGCAGAACCTGAGCTGCTGCTGCTGCTGCTG 1372
Db 1153 CGTGCGGCTCTTTAGGCGAGCTGCAAAATAGGCAGAACCTGAGCTGCTGCTGCTGCTGCTG 1212

QY 1373 TCTGACAGACACGCTGAGTGTGATGAAAGTGTGAGGCTGCTGCTGCTGCTGCTGCTGCTG 1432
Db 1213 TCTGACAGACACGCTGAGTGTGATGAAAGTGTGAGGCTGCTGCTGCTGCTGCTGCTGCTG 1272

QY 1433 GCTGTGCGGAGACGATTTGAGAGCTGAGAGAAACCTGCTGCTGCTGCTGCTGCTGCTGCT 1492
Db 1273 GCTGTGCGGAGACGATTTGAGAGCTGAGAGAAACCTGCTGCTGCTGCTGCTGCTGCTGCT 1332

QY 1493 GAGAGCAGATGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 1552
Db 1333 GAGAGCAGATGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 1392

QY 1553 CAGAAATG 1560
Db 1393 CAGAAATG 1400

RESULT 11
E43450 969 bp DNA linear PAT 31-JAN-2002
LOCUS Novel protein G-coupled receptor protein and DNA thereof.
DEFINITION
ACCESSION E43450
VERSION E43450.1 GI:18627716
KEYWORDS JP 200016576-A/1.
SOURCE Homo sapiens (human)
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Homidae; Homo.

REFERENCE 1 (bases 1 to 969)
AUTHORS Matanabe, T., Terao, S. and Matsui, H.
TITLE Novel protease G-coupled receptor protein and DNA thereof
JOURNAL Patent: JP 2000166576-A 1 20-JUN-2000;
TAKEDA CHEM IND LTD
OS Homo sapiens (human)
PN JP 2000166576-A/1
PD 20-JUN-2000
PR 30-SEP-1999 JP 1999278116
PR
PI TAKUYA WATANABE, SHIZUKO TERAO, HIDEKI MATSUI
PC C12N15/00, A61K38/00, A61K39/395, A61K45/00, A61K48/00,
PC A61P43/00,
PC C07K14/705, C07K16/28, C12N1/21, C12N5/10, C12P21/02, C12P21/08, PC
G01N33/15,
PC G01N33/53, G01N33/566, G01N33/577// (C12N1/21, C12R1:19), (C12N5/10, PC
C12R1:91),
PC (C12P21/02, C12R1:19), (C12P21/08, C12R1:91), C12N15/00, A61K37/02,
PC C12N5/00, C12R1:91)
PC (C12N5/00, C12R1:91)
CC
FH
FT
FT
Key Location/Qualifiers
source 1. 969
/organism="Homo sapiens"
/mol_type="genomic DNA"
/db_xref="taxon:9606"
ORIGIN
Query Match 37.0%; Score 969; DB 6; Length 969;
Best Local Similarity 100.0%; Pred. No. 0;
Matches 969; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 492 ATGATTCACCAACCCAGTCTTGGGTACAGAACTGACCAACATCAACGAGCGAGAG 551
DB 1 ATGATTCACCAACCCAGTCTTGGGTACAGAACTGACCAACATCAACGAGCGAGAG 60
QY 552 ACTCTTGTCTACAGAGAACCCCTGAGCTTCAACGGGGCTGACGTGATGTTCCCTTGTG 611
DB 61 ACTCTTGTCTACAGAGAACCCCTGAGCTTCAACGGGGCTGACGTGATGTTCCCTTGTG 120
QY 612 GCGCTGACAGAGAACCGGGTGTGCTCTGAGCTCTGAGGCTCCCGATCGCAGAGAACCT 671
DB 121 GCGCTGACAGAGAACCGGGTGTGCTCTGAGCTCTGAGGCTCCCGATCGCAGAGAACCT 180
QY 672 GTCTCCATCTACATCTTCAACCTGAGTGGGGCCGACCTTCTCTTCTTACGGGCGACATT 731
DB 181 GTCTCCATCTACATCTTCAACCTGAGTGGGGCCGACCTTCTCTTCTTACGGGCGACATT 240
QY 732 ATATGTTGCGCGTTACGCTCATCATATATCCGCCATCCATCTCCAAATCTTCAGTCT 791
DB 241 ATATGTTGCGCGTTACGCTCATCATATATCCGCCATCCATCTCCAAATCTTCAGTCT 300
QY 792 GTATGACCTTTCCCTACTTATAGGCTTAAGCATCTGAGCGCATGACACCGAGCGC 851
DB 301 GTATGACCTTTCCCTACTTATAGGCTTAAGCATCTGAGCGCATGACACCGAGCGC 360
QY 852 TGCGCTGACCTCTGAGCGCATCTGAGTACATGCGCGCGCGCCGACGATCTGTCATCG 911
DB 361 TGCGCTGACCTCTGAGCGCATCTGAGTACATGCGCGCGCGCCGACGATCTGTCATCG 420
QY 912 GTATGATGTCTCTGCTCTGAGCGCTTCTGCTGCTGAGTACATCTGAGTGTTC 971
DB 421 GTATGATGTCTCTGCTCTGAGCGCTTCTGCTGCTGAGTACATCTGAGTGTTC 480
QY 972 TGATGATCTCTGTTAGTGTGCTGATCTGTTTGTGTGAAAGCTGATTTCTTACA 1031
DB 481 TGATGATCTCTGTTAGTGTGCTGATCTGTTTGTGTGAAAGCTGATTTCTTACA 540
QY 1032 ATGCGGTGCGTGTGTTTATATGTGTGCTCTGCTGAGGCTGACGCTGCTGTGCTC 1091

DB 541 ATGCGGTGCGTGTGTTTATATGTGTGCTCTGCTGAGGCTGACGCTGCTGTGCTC 600
QY 1092 AGATTTCTCTGTGATCCCGGAAAGATGCGGCTGACAGAGCTGTATGATCAATCTCTC 1151
DB 601 AGATTTCTCTGTGATCCCGGAAAGATGCGGCTGACAGAGCTGTATGATCAATCTCTC 660
QY 1152 ACAGTGTGCTCTCTCTCTGCTGAGCGCTTCTGCTGCTTGTGATTCAGTGTGCTTCTC 1211
DB 661 ACAGTGTGCTCTCTCTCTGCTGAGCGCTTCTGCTGCTTGTGATTCAGTGTGCTTCTC 720
QY 1212 AGATTCACCTGATGTAAGTGTATTTTGTGATGATGATGATGATTTTCTG 1271
DB 721 AGATTCACCTGATGTAAGTGTATTTTGTGATGATGATGATGATTTTCTG 780
QY 1272 TCCGCTCTTACAGAGTGTGCAACCCCATCTTACTTCTCTGAGGCTCTTAAAGGAG 1331
DB 781 TCCGCTCTTACAGAGTGTGCAACCCCATCTTACTTCTCTGAGGCTCTTAAAGGAG 840
QY 1332 CGTCAAAATAGAGAGAACCTGAGCTGCTTCCAGAGAGGCTTGCAGAGACAGCGCTGAG 1391
DB 841 CGTCAAAATAGAGAGAACCTGAGCTGCTTCCAGAGAGGCTTGCAGAGACAGCGCTGAG 900
QY 1392 GTGATGAGGTGAGAGGCTGCTTCTTCAAGAAACCTGAGCTGTGCGAAGAGATTG 1451
DB 901 GTGATGAGGTGAGAGGCTGCTTCTTCAAGAAACCTGAGCTGTGCGAAGAGATTG 960
QY 1452 GAGCAGTGA 1460
DB 961 GAGCAGTGA 969
RESULT 12
AX498196
LOCUS AX498196 969 bp DNA linear PAT 26-SEP-2002
DEFINITION Sequence 19 from Patent WO242461.
ACCESSION AX498196
VERSION AX498196.1 GI:23343119
KEYWORDS
SOURCE
ORGANISM
synthetic construct
synthetic construct
artificial sequences.
REFERENCE
1
AUTHORS Chen, R., Chu, Z. L., Dang, H. T., Lowitz, K. P. and Pride, C.
TITLE Endogenous and non-endogenous versions of human g protein-coupled
JOURNAL Patent: WO 0242461-A 19 30-MAY-2002;
Arena Pharmaceuticals, Inc. (US)
FEATURES
source
1. 969
Location/Qualifiers
/organism="synthetic construct"
/mol_type="unassigned DNA"
/db_xref="taxon:32630"
/note="Novel Sequence"
ORIGIN
Query Match 37.0%; Score 969; DB 6; Length 969;
Best Local Similarity 100.0%; Pred. No. 0;
Matches 969; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 492 ATGATTCACCAACCCAGTCTTGGGTACAGAACTGACCAACATCAACGAGCGAGAG 551
DB 1 ATGATTCACCAACCCAGTCTTGGGTACAGAACTGACCAACATCAACGAGCGAGAG 60
QY 552 ACTCTTGTCTACAGAGAACCCCTGAGCTTCAACGGGGCTGACGTGATGTTCCCTTGTG 611
DB 61 ACTCTTGTCTACAGAGAACCCCTGAGCTTCAACGGGGCTGACGTGATGTTCCCTTGTG 120
QY 612 GCGCTGACAGAGAACCGGGTGTGCTCTGAGCTCTGAGGCTCCCGATCGCAGAGAACCT 671
DB 121 GCGCTGACAGAGAACCGGGTGTGCTCTGAGCTCTGAGGCTCCCGATCGCAGAGAACCT 180
QY 672 GTCTCCATCTACATCTTCAACCTGAGTGGGGCCGACCTTCTCTTCTTACGGGCGACATT 731

AUTHORS Lembo, P.M., Graziani, E., Groblewski, T., O'Donnell, D., Roy, M.-O., Zhang, J., Hoffert, C., Cao, D., Schmidt, R., Pelletier, M., Labarre, M., Gosselin, M., Fortin, Y., Banville, D., Shen, S., Strom, P., Payza, K., Dray, A., Walker, P. and Ahmad, S.

TITLE Premecephalin A gene products activate a new family of sensory neuron-specific GPCRs

JOURNAL Nat. Neurosci. 5 (3), 201-209 (2002).

REFERENCE 2 (bases 1 to 969)

PubMed 11850634

ALTHORS Ahmad, S., Banville, D., Fortin, Y., Lembo, P.M. and O'Donnell, D.

TITL Direct Submission

JOURNAL Submitted (25-JAN-2002) Molecular Biology, AstraZeneca, 7171 Frederick-Banting, Montreal, Quebec H4S 1Z9, Canada

COMMENT On Jun 16, 2003 this sequence version replaced gi:19338907.

FEATURES location/Qualifiers

source 1..969

organism "Homo sapiens"

mol_type "genomic DNA"

db_xref "taxon:9606"

product "G protein-coupled receptor SNR1"

note "GPCR"

codon_start 1

product "G protein-coupled receptor SNR1"

protein_id "AAL6878.2"

db_xref "gi:31747895"

translation "MDPTIPVCTELTPINGREBPYKQTLSPFGLNCTIVSLVLTGNAVIMLIGCRKRRRAVSIYIINLVADPLFSGHIGPLINIRHPSIKLSPVMTFPIYIGLSMLSAISIERCLSLMPWICRPRYSVYCVLWMLSLRILIMWFCDPLFSGANSVWCETSDPILAVFLCVLGGSSIVLVRLICSRKMLRLVYVILLVAVFLCGLPFGIOWALFSRILHDMVLECHVALVSIFFSLANSANPPIYFVGSFRQRNQNKLVLVGRALQDPPEVDEGGMLPQETLELSGKLEQ"

ORIGIN

Query Match 30.7%; Score 805; DB 9; Length 969;
 Best Local Similarity 99.8%; Pred. No. 0;
 Matches 905; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

499 CAACCATCCGAGTCTGGTGTGAGACAGAACTGACCAATCAACGAGCTGAGAGACCTCTT 558
 |||||
 8 CAACCATCCGAGTCTGGTGTGAGACAGAACTGACCAATCAACGAGCTGAGAGACCTCTT 67
 |||||
 559 GCTACAAAGCAGACCTGAGCTTCAAGGAGCTGACGTCATCGTTCCCTGTGCGCTGA 618
 |||||
 68 GCTACAAAGCAGACCTGAGCTTCAAGGAGCTGACGTCATCGTTCCCTGTGCGCTGA 127
 |||||
 619 CAGAAACGCGGTGTGCTCTGAGCTTCTGAGGCTGCGCATGCGCAGAAACGCTGTCTCA 678
 |||||
 128 CAGAAACGCGGTGTGCTCTGAGCTTCTGAGGCTGCGCATGCGCAGAAACGCTGTCTCA 187
 |||||
 679 TCTACATCTGAACTGAGTGGCGGCGCACTTCTCTTCTAGCGGCGCAATTATATGTT 738
 |||||
 188 TCTACATCTGAACTGAGTGGCGGCGCACTTCTCTTCTAGCGGCGCAATTATATGTT 247
 |||||
 739 CGCGTTAGCGCTCATCAATATCCGCAATCCATCTCAAAATCTCACTCTGTGATGA 798
 |||||
 248 CGCGTTAGCGCTCATCAATATCCGCAATCCATCTCAAAATCTCACTCTGTGATGA 307
 |||||
 799 CTTTCCCTACTTATAGGCTTAAGCCTAGAGCGCATCAGACCGAGCGCTGCTGT 858
 |||||
 308 CTTTCCCTACTTATAGGCTTAAGCCTAGAGCGCATCAGACCGAGCGCTGCTGT 367
 |||||
 859 CCATCTGTGGCGCATCTGTGACCACTGCGCGCGCGCGCGCATCTGTATCTGTATGT 918
 |||||
 368 CCATCTGTGGCGCATCTGTGACCACTGCGCGCGCGCGCGCATCTGTATCTGTATGT 427
 |||||
 919 GTGTCTGTCTCTGGAGCGCTGTCCCTGTCTCTGAGAGTATCTGTGAGTGTCTGTGACT 978
 |||||
 428 GTGTCTGTCTCTGGAGCGCTGTCCCTGTCTCTGAGAGTATCTGTGAGTGTCTGTGACT 487
 |||||
 979 TCTGTGTAGTGTGCTGATCTGTGTGTGTGAAACGTCAATTTCATTACATGCGGT 1038
 |||||

|||||
 488 TCTGTGTAGTGTGCTGATCTGTGTGTGTGAAACGTCAATTTCATTACATGCGGT 547
 |||||
 1039 GCGT 1098
 |||||
 548 GCGT 607
 |||||
 1099 TCTGTGTATCCCGAAGATGCCGTGACCAAGCTGTACATCACTCTCTCAAGTGC 1158
 |||||
 608 TCTGTGTATCCCGAAGATGCCGTGACCAAGCTGTACATCACTCTCTCAAGTGC 667
 |||||
 1159 TGTCTCTCTCTCTGT 1218
 |||||
 668 TGTCTCTCTCTCTGT 727
 |||||
 1219 ACCGTGATTTGAAAGCTTATTTGTGATGTGATGTGATGTGATGTGATGTGATGTGAT 1278
 |||||
 728 ACTGTGATTTGAAAGCTTATTTGTGATGTGATGTGATGTGATGTGATGTGATGTGAT 787
 |||||
 1279 TTAACAGACGTGCCAACCCCATCTTACTTCTGTGTGTGTGTGTGTGTGTGTGTGT 1338
 |||||
 788 TTAACAGACGTGCCAACCCCATCTTACTTCTGTGTGTGTGTGTGTGTGTGTGTGT 847
 |||||
 1339 ATAGCAGAACTGAAAGCTGTCTCCAGAGGCTCTGACAGACACGCTGAGGTGATG 1398
 |||||
 848 ATAGCAGAACTGAAAGCTGTCTCCAGAGGCTCTGACAGACACGCTGAGGTGATG 907
 |||||
 1399 AAGGTG 1405
 |||||
 908 AAGGTG 914
 |||||

Search completed: October 29, 2004, 22:34:19
 Job time : 10845 secs

GenCore version 5.1.6
Copyright (c) 1993 - 2004 Compugen Ltd.

OM nucleic - nucleic search, using sw model

Run on: October 29, 2004, 16:07:33 ; Search time 1225 Seconds
(without alignments)
11218.763 Million cell updates/sec

Title: US-09-867-570-1

Perfect score: 2618
Sequence: 1 aacaatgcgcgcgaattcgg.....gaaaaaaaaaaaaaaaaa 2618

Scoring table: OLIGO_NUC
Gapop 60.0 , Gapext 60.0

Searched: 4134886 seqs, 2624710521 residues

Word size : 20

Total number of hits satisfying chosen parameters: 41862

Minimum DB seq length: 0

Maximum DB seq length: 200000000

Post-processing: Listing first 45 summaries

Database :

N_Geneseq_23Sep04:*
1: geneseqn1980s:*
2: geneseqn1980s:*
3: geneseqn2000s:*
4: geneseqn2001as:*
5: geneseqn2001bs:*
6: geneseqn2002as:*
7: geneseqn2002bs:*
8: geneseqn2003as:*
9: geneseqn2003bs:*
10: geneseqn2003cs:*
11: geneseqn2003ds:*
12: geneseqn2004s:*

Pred. No. is the number of results predicted by chance to have a
score greater than or equal to the score of the result being printed,
and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length DB	ID	Description
1	2618	100.0	2618	6	ABK52822
2	1101	42.1	6	ABK52823	ADK52822 CDNA enco
3	1093	41.7	1369	10	ADK52823 Genomic D
4	1088	41.6	1400	8	ADK52823 Human GPC
5	1088	41.6	1400	8	ADK52823 Human GPC
6	1088	41.6	1400	8	ADK52823 Human GPC
7	969	37.0	969	3	AAA29811
8	969	37.0	969	3	AAA29811
9	969	37.0	969	3	AAA29811
10	969	37.0	969	3	AAA29811
11	969	37.0	969	3	AAA29811
12	969	37.0	969	3	AAA29811
13	969	37.0	969	3	AAA29811
14	969	37.0	969	3	AAA29811
15	969	37.0	969	3	AAA29811
16	969	37.0	969	3	AAA29811
17	969	37.0	969	3	AAA29811
18	969	37.0	969	3	AAA29811
19	969	37.0	969	3	AAA29811
20	969	37.0	969	3	AAA29811
21	969	37.0	969	3	AAA29811

22	299	11.4	954	10	ADK12723	Adc12723 Human GPC
23	291	11.1	291	8	ACA04716	ACA04716 CDNA enco
24	275	10.5	275	10	ACA55732	ACA55732 Human sig
25	275	10.5	275	12	AD155528	Ad155528 Human pol
26	270	10.3	1323	5	AA858260	AA858260 DNA encod
27	179	6.8	969	2	AA210069	AA210069 Human dor
28	120	4.6	964	6	AD116637	Ad116637 Human NOV
29	120	4.6	964	12	ADN42291	Adn42291 Human cdn
30	120	4.6	969	2	AA210070	AA210070 Human dor
31	120	4.6	969	2	AA209943	AA209943 Human G P
32	120	4.6	969	5	AA507944	AA507944 Human cdn
33	120	4.6	969	6	AA274949	AA274949 Human G-P
34	120	4.6	969	6	AB242891	AB242891 Human GPC
35	120	4.6	969	8	ACCA4772	ACCA4772 Human G P
36	120	4.6	969	8	AB281334	AB281334 Human dor
37	120	4.6	969	10	AA61647	AA61647 Human GPC
38	120	4.6	969	11	AD196465	AD196465 Human G P
39	120	4.6	969	12	AD030078	AD030078 Human GPC
40	120	4.6	973	4	AAK51462	AAK51462 Human pol
41	120	4.6	997	6	AD116635	AD116635 Human NOV
42	120	4.6	997	12	ADN42289	ADN42289 Human cdn
43	120	4.6	1116	4	AAH49507	AAH49507 Human GTP
44	120	4.6	1369	10	ADK6444	ADK6444 Human GPC
45	120	4.6	2040	6	AA23744	AA23744 Human Mrg

ALIGNMENTS

RESULT 1	ABK52822	ABK52822 standard; CDNA; 2618 BP.
ID	ABK52822	
XX	ABK52822	
AC	ABK52822	
DT	27-AUG-2002	(first entry)
XX		
DE	CDNA encoding human G-protein coupled receptor (GPCR).	
XX		
KW	Human; G-protein coupled; receptor; GPCR; human protease;	
KW	human therapeutic protein; query sequence; search; gene; seq;	
KW	sequence database; non-human transgenic animal; gene therapy;	
KW	chromosome 3.	
XX		
OS	Homo sapiens.	
XX		
FH	Key	Location/Qualifiers
FT	CDS	447..1460
FT		/*tag= a
FT		/product= "Human G-protein coupled receptor (GPCR) "
XX		
PN	W0200234914-A1.	
XX		
PD	02-MAY-2002.	
XX		
PF	10-OCT-2001; 2001MO-US031592.	
XX		
PR	25-OCT-2000; 2000US-00695045.	
PR	31-MAY-2001; 2001US-00867570.	
XX		
PA	(PEKE) PE CORP NY.	
XX		
FI	Wei M, Zhao Q, Woodage T, Di Francesco V, Beasley EW;	
XX		
DR	WPI; 2002-463360/49.	
XX		
DR	P-PSDB; AAU97598.	
XX		
PT	Novel isolated G-protein coupled receptor peptide useful for treating	
PT	disorder characterized by absence of, in appropriate or unwanted	
PT	expression of the receptor protein, and as immunogens to raise	
XX	antibodies.	
PS	Claim 4; Fig 1; 75bp; English.	

QY 1861 TCATGTTCCCTTTATGACTGAGGCACTTACTGCACTGGAGAGCTCAATCTTAATAGT 1920
 DB 1861 TCATGTTCCCTTTATGACTGAGGCACTTACTGCACTGGAGAGCTCAATCTTAATAGT 1920
 QY 1921 GAGTTCGTCTACTCTTAATTCCTGATTTCTGATATTAAGCAAAATATGACCTTA 1980
 DB 1921 GAGTTCGTCTACTCTTAATTCCTGATTTCTGATATTAAGCAAAATATGACCTTA 1980
 QY 1981 GAGGAGATCTCCCTCATTAATAAGCTCTGATTAATGCTTTATGAATAGCCCTCTC 2040
 DB 1981 GAGGAGATCTCCCTCATTAATAAGCTCTGATTAATGCTTTATGAATAGCCCTCTC 2040
 QY 2041 CTGTCAATTTGTCACACAGCTGTCACATGTTGGCTTGTCTTCTAGTAAGACATGCTG 2100
 DB 2041 CTGTCAATTTGTCACACAGCTGTCACATGTTGGCTTGTCTTCTAGTAAGACATGCTG 2100
 QY 2101 GCCCTTCCCTTGAGAGCTGTAAGTCTTATTTAGCTCTTCTGACATTAATGACTAG 2160
 DB 2101 GCCCTTCCCTTGAGAGCTGTAAGTCTTATTTAGCTCTTCTGACATTAATGACTAG 2160
 QY 2161 TGAGAGCTTAAATATGTCACACAGCTTCAATTTGGCATTGGAACCTCAATATG 2220
 DB 2161 TGAGAGCTTAAATATGTCACACAGCTTCAATTTGGCATTGGAACCTCAATATG 2220
 QY 2221 ATTTAAAGTGAATATATCTTGAACCAATTTATTAATCACTACAGATTTCTTCACTT 2280
 DB 2221 ATTTAAAGTGAATATATCTTGAACCAATTTATTAATCACTACAGATTTCTTCACTT 2280
 QY 2281 GTAGAGAAATCTTCACTTCCAGGTTTGTATTAATTTGTTCTGATTTGTAATCTTCACT 2340
 DB 2281 GTAGAGAAATCTTCACTTCCAGGTTTGTATTAATTTGTTCTGATTTGTAATCTTCACT 2340
 QY 2341 TAGTTTAAAGCTTGTATTAATGAGCAAACTGAAACATCTGACCTTCCATGCA 2400
 DB 2341 TAGTTTAAAGCTTGTATTAATGAGCAAACTGAAACATCTGACCTTCCATGCA 2400
 QY 2401 TCTCAATATGATCTGATTAATTAATCACTGATTTGTAACAGATTTCTGATCACTGCT 2460
 DB 2401 TCTCAATATGATCTGATTAATTAATCACTGATTTGTAACAGATTTCTGATCACTGCT 2460
 QY 2461 GACATCATGATCACTGGAATATTTGCTTAAGGAAATTAAGCCAGACCAACATATTT 2520
 DB 2461 GACATCATGATCACTGGAATATTTGCTTAAGGAAATTAAGCCAGACCAACATATTT 2520
 QY 2521 GTAAAGTCAATCTTAAGGATATCCAAATTAAGAAATTTCTGAAACAGAAATTAATTT 2580
 DB 2521 GTAAAGTCAATCTTAAGGATATCCAAATTAAGAAATTTCTGAAACAGAAATTAATTT 2580
 QY 2581 AGGAGATCTGCTGCTGGAATTAATTTGCTTAAGGAAATTAAGCCAGACCAACATATTT 2640
 DB 2581 AGGAGATCTGCTGCTGGAATTAATTTGCTTAAGGAAATTAAGCCAGACCAACATATTT 2640
 RESULT 2
 ID ABK52823 standard; DNA; 8622 BP.
 AC ABK52823;
 DT 27-AUG-2002 (first entry)
 DE Genomic DNA encoding human G-protein coupled receptor (GPCR).
 XX Human; G-protein coupled; receptor; GPCR; human protease;
 XX human therapeutic protein; query sequence; search; gene; ds;
 KW sequence database; non-human transgenic animal; gene therapy;
 XX chromosome 3.
 OS Homo sapiens.
 XX Key Location/Qualifiers
 FH CDS 4300..8496
 FT

FT /*tag= a
 FT /product= "Human G-protein coupled receptor (GPCR)"
 FT exon 4300..4319
 FT /*tag= b
 FT /number= 1
 FT intron 4320..7502
 FT /*tag= c
 FT /number= 1
 FT exon 7503..8496
 FT /*tag= d
 FT /number= 2
 PN WO200234914-A1.
 PD 02-MAY-2002.
 XX 10-OCT-2001; 2001NO-US031592.
 XX 25-OCT-2000; 2000US-00695045.
 PR 31-MAY-2001; 2001US-00867570.
 XX
 PA (PEKE) PE CORP NY.
 XX Wei M, Zhao Q, Woodage T, Di Francesco V, Beasley EM;
 PI WPI; 2002-463360/49.
 DR P-PSDB; AAU97598.
 XX
 PT Novel isolated G-protein coupled receptor peptide useful for treating
 PT disorder characterized by absence of, in appropriate or unwanted
 PT expression of the receptor protein, and as immunogens to raise
 PT antibodies.
 XX
 PS Claim 4; Fig 3; 75pp; English.
 XX
 CC The present invention relates to a new G-protein coupled receptor (GPCR)
 CC peptide. The invention is useful for identifying a modulator of GPCR and
 CC for treating a disease or condition mediated by a human protease. The
 CC invention is also useful as models for the development of human
 CC therapeutic, for identifying therapeutic proteins, as targets for
 CC development of human therapeutic agents, and as query sequence to perform
 CC a search against sequence databases to, for e.g., identify other family
 CC members of related sequences. The vector of the invention is useful for
 CC producing a GPCR protein or peptide, for conducting cell-based assays
 CC involving the GPCR protein or its fragment, for identifying GPCR protein
 CC mutants whose functions are affected, and to produce non-human transgenic
 CC animals. The present nucleic acid sequence represents the human G-protein
 CC coupled receptor (GPCR) gene located on chromosome 3. This sequence
 CC encodes the human G-protein coupled receptor (GPCR) protein of the
 CC invention
 CC
 SQ Sequence 8622 BP; 2684 A; 1921 C; 1679 G; 2338 T; 0 U; 0 Other:
 Query Match 42.1%; Score 1101; DB 6; Length 8622;
 Best Local Similarity 100.0%; Pred. No. 0;
 Matches 1101; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
 QY 466 GGGTCATCAGACTGGGGTTTCTGAGCTGATTCACCATCCAGTTGGGTAACAGAC 525
 DB 7502 GGGTCATCAGACTGGGGTTTCTGAGCTGATTCACCATCCAGTTGGGTAACAGAC 7561
 QY 526 TGACACCAATCAACGAGCGTGAAGAGATCTCTGTAAGAGAGACCTGAGCTTCAAG 585
 DB 7562 TGACACCAATCAACGAGCGTGAAGAGATCTCTGTAAGAGAGACCTGAGCTTCAAG 7621
 QY 586 GGGTCAGTGCATCGTTTCCCTTGTGCGGCTGACAGGAAAGCGGTTGTCTGTGCTCC 645
 DB 7622 GGGTCAGTGCATCGTTTCCCTTGTGCGGCTGACAGGAAAGCGGTTGTCTGTGCTCC 7681
 QY 646 TGGGCTGCCGATGCGGAGAGAGCGTGTCTCACTTAACCTCAACCTGTGGCGGCGG 705
 DB 7682 TGGGCTGCCGATGCGGAGAGAGCGTGTCTCACTTAACCTCAACCTGTGGCGGCGG 7741

QY 706 ACTTCCTCTCTTACGGCCGACATTATATGTTGCGCTTACGCTCATCAATATCCGCC 765
DB 7742 ACTTCCTCTCTTACGGCCGACATTATATGTTGCGCTTACGCTCATCAATATCCGCC 7801
QY 766 ATCCCATCTCCAAATCTCAGTCCGTGATGACCTTTCCACTTATATAGGCTTAAGCA 825
DB 7802 ATCCCATCTCCAAATCTCAGTCCGTGATGACCTTTCCACTTATATAGGCTTAAGCA 7861
QY 826 TGCTGAGGCGCATGACGACCCGAGCGCTGCTGTCATCTGATGAGCCCATCTGATACACT 885
DB 7862 TGCTGAGGCGCATGACGACCCGAGCGCTGCTGTCATCTGATGAGCCCATCTGATACACT 7921
QY 886 GCGCGCGCCGACGATACCTGTCATGTCATGTCGTCCTGCTGAGGCGCTGCTGCTG 945
DB 7922 GCGCGCGCCGACGATACCTGTCATGTCATGTCGTCCTGCTGAGGCGCTGCTGCTG 7981
QY 946 TGGGAGATATCTGAGTGGATGATGTTCTGATCTCTGTTTATGAGTGGTGTGATTCGTT 1005
DB 7982 TGGGAGATATCTGAGTGGATGATGTTCTGATCTCTGTTTATGAGTGGTGTGATTCGTT 8041
QY 1006 GGTGTGAACGTGACGATTTTCAATACATCGCGTGGTGGTGGTGGTGGTGGTGGTGGT 1065
DB 8042 GGTGTGAACGTGACGATTTTCAATACATCGCGTGGTGGTGGTGGTGGTGGTGGTGGT 8101
QY 1066 GTGGGTCAGCGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 1125
DB 8102 GTGGGTCAGCGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 8161
QY 1126 CCAGGCTGTAGTGACGATCCCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 1185
DB 8162 CCAGGCTGTAGTGACGATCCCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 8221
QY 1186 TTGGCATTCAGTGGGCGCTGTTTTCAGAGATCCACTGATGTAAGTCTTATTTTGTG 1245
DB 8222 TTGGCATTCAGTGGGCGCTGTTTTCAGAGATCCACTGATGTAAGTCTTATTTTGTG 8281
QY 1246 ATGTGATCTATGTTTCCATTTTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 1305
DB 8282 ATGTGATCTATGTTTCCATTTTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 8341
QY 1306 ACTTCCTCTCTTACGGCCGACATTATATGTTGCGCTTACGCTCATCAATATCCGCC 1365
DB 8342 ACTTCCTCTCTTACGGCCGACATTATATGTTGCGCTTACGCTCATCAATATCCGCC 8401
QY 1366 AGAGGCGCTCTGACGACGCGCTGAGTGAATGAAGTGAAGTGAAGTGAAGTGAAGTGA 1425
DB 8402 AGAGGCGCTCTGACGACGCGCTGAGTGAATGAAGTGAAGTGAAGTGAAGTGAAGTGA 8461
QY 1426 CCCTGAGAGCTGTGCGGAAAGCAGATTGAGCAGTGAAGAACTCTGCTGCTGCTGCTG 1485
DB 8462 CCCTGAGAGCTGTGCGGAAAGCAGATTGAGCAGTGAAGAACTCTGCTGCTGCTGCTG 8521
QY 1486 GAGCTTTGAGAGCATATGTCGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 1545
DB 8522 GAGCTTTGAGAGCATATGTCGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 8581
QY 1546 TCTGCTCAGAAATGCTCAG 1566
DB 8582 TCTGCTCAGAAATGCTCAG 8602

RESULT 3
ID ADC86820 standard; DNA; 1369 BP.
XX AC ADC86820;
XX 01-JAN-2004 (first entry)
XX Human GPCR gene SEQ ID NO:1273.
XX de; gene; human; GPCR;
KW guanosine triphosphate-binding protein coupled receptor; gene therapy.

XX OS Homo sapiens.
XX PN EP1270724-A2.
XX 02-JAN-2003.
XX 18-JUN-2002; 2002EP-00013517.
XX 18-JUN-2001; 2001JP-00246789.
XX (MAD-) NAT INST ADVANCED IND SCI & TECHNOLOGY.
XX (ADSC-) CENT ADVANCED SCI & TECHNOLOGY INCUBARIO.
XX Suwa M, Arai K, Akiyama Y, Aburatani H;
XX MPI; 2003-115783/31.
XX P-PsDB; ADC86821.
XX New polynucleotide, useful for preparing a composition for treating a
XX patient in need of increased or suppressed activity or expression of the
XX guanosine triphosphate-binding protein coupled receptor.
XX Claim 1; SEQ ID NO 1273; 28bp; English.
XX The invention relates to a novel polynucleotide encoding a guanosine
XX triphosphate-binding protein coupled receptor (GPCR). A polynucleotide of
XX the invention may have a use in gene therapy. The polynucleotide and
XX polypeptide are useful for preparing a composition for treating a patient
XX in need of increased or suppressed activity or expression of the
XX guanosine triphosphate-binding protein coupled receptor. The
XX polynucleotide sequences shown in ADC85548-ADC87616 encode GPCR's of the
XX invention.
SQ Sequence 1369 BP; 272 A; 372 C; 340 G; 385 T; 0 U; 0 Other;
Query Match 41.7%; Score 1093; DB 10; Length 1369;
Best Local Similarity 99.8%; Pred. No. 0;
Matches 1193; Conservative 0; Mismatches 2; Indels 0; Gaps 0;
QY 466 GGGTATCATGAGTGGGAGTTTCTGAGCATGATTAACCATCCAGTCTTGGGTACAGAAC 525
DB 175 GGGTATCATGAGTGGGAGTTTCTGAGCATGATTAACCATCCAGTCTTGGGTACAGAAC 234
QY 526 TGACACCAATCAACGAGAGTGAAGAGTCTCTGTCACAGACACCTGAGCTTCAAGG 585
DB 235 TGACACCAATCAACGAGAGTGAAGAGTCTCTGTCACAGACACCTGAGCTTCAAGG 294
QY 586 GAGTGAAGTGAATGCTTTTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 645
DB 295 GAGTGAAGTGAATGCTTTTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 354
QY 646 TGGGCTGCGCGATGCGGAGAAAGCTGCTCATCTCAATCTCAACCTGCTGCGGCGG 705
DB 355 TGGGCTGCGCGATGCGGAGAAAGCTGCTCATCTCAATCTCAACCTGCTGCGGCGG 414
QY 706 ACTTCCTCTCTTACGGCCGACATTATATGTTGCGCTTACGCTCATCAATATCCGCC 765
DB 415 ACTTCCTCTCTTACGGCCGACATTATATGTTGCGCTTACGCTCATCAATATCCGCC 474
QY 766 ATCCCATCTCCAAATCTCAGTCCGTGATGACCTTTCCACTTATATAGGCTTAAGCA 825
DB 475 ATCCCATCTCCAAATCTCAGTCCGTGATGACCTTTCCACTTATATAGGCTTAAGCA 534
QY 826 TGCTGAGGCGCATGACGACCCGAGCGCTGCTGTCATCTGATGAGCCCATCTGATACACT 885
DB 535 TGCTGAGGCGCATGACGACCCGAGCGCTGCTGTCATCTGATGAGCCCATCTGATACACT 594
QY 886 GCGCGCGCCGACGATACCTGTCATGTCATGTCGTCCTGCTGAGGCGCTGCTGCTGCTG 945
DB 595 GCGCGCGCCGACGATACCTGTCATGTCATGTCGTCCTGCTGAGGCGCTGCTGCTGCTG 654
QY 946 TGGGAGATATCTGAGTGGATGATGTTCTGATCTCTGTTTATGAGTGGTGTGATTCGTT 1005

```
Db      655 TCGGAGATATCTGAGTGTGATGTTCTGTGATCTTCCGTTAGTGGTCAATATCTGTTT 714
Qy      1006 GGGTGAAGAGTGAATTTGATTAACATTCGGTGGCTGGTTTTTATATGTTGGTCTCT 1065
Db      715 GGTGTGAAGAGTGAATTTGATTAACATTCGGTGGCTGGTTTTTATATGTTGGTCTCT 774
Qy      1066 GTGGGTTCAGACCTGTGTCCTGTGTCAGAGATTTCTCTGTGATCCCGAAGATGCGCTGA 1125
Db      775 GTGGGTTCAGACCTGTGTCCTGTGTCAGAGATTTCTCTGTGATCCCGAAGATGCGCTGA 834
Qy      1126 CCAAGGCTGATGATGATCCATCTCTCTCAAGTGTGTTCTTCTCTGTGGCTGCTCT 1185
Db      835 CCAAGGCTGATGATGATCCATCTCTCTCAAGTGTGTTCTTCTCTGTGGCTGCTCT 894
Qy      1186 TTGGCATATGATGGGCTGTTTTCAGATTCACCTGATGATGGAAGTCTATTTTGTCT 1245
Db      895 TTGGCATATGATGGGCTGTTTTCAGATTCACCTGATGATGGAAGTCTATTTTGTCT 954
Qy      1246 ATGTCATCTAGTTTTCATTTTCTGTCGCTCTTAAACAGATGCGCAACCCCATCATTT 1305
Db      955 ATGTCATCTAGTTTTCATTTTCTGTCGCTCTTAAACAGATGCGCAACCCCATCATTT 1014
Qy      1306 ACTTCTTCTGTCGCTCTTAAAGCAGCTCAAAATAGCAGAACCTGAGCTGTTCTCC 1365
Db      1015 ACTTCTTCTGTCGCTCTTAAAGCAGCTCAAAATAGCAGAACCTGAGCTGTTCTCC 1074
Qy      1366 AGAGGGCTCTGACAGACACCTGATGATGATGATGATGATGATGATGATGATGATGAT 1425
Db      1075 AGAGGGCTCTGACAGACACCTGATGATGATGATGATGATGATGATGATGATGATGAT 1134
Qy      1426 CCTGAGAGCTGTGGGAGAGAGATTTGAGAGATGAGAGAGAGAGAGAGAGAGAGAGAGAG 1485
Db      1135 CCTGAGAGCTGTGGGAGAGAGATTTGAGAGATGAGAGAGAGAGAGAGAGAGAGAGAGAG 1194
Qy      1486 GAACTTTGAGAGAGATGTCGCTGCAACCTTGTGACATTAATATGATTTTCTTACCT 1545
Db      1195 GAACTTTGAGAGAGATGTCGCTGCAACCTTGTGACATTAATATGATTTTCTTACCT 1254
Qy      1546 TCTGCTCTGAGAAATGTCATGATGTCCTCAAGGCTTTCGATATGATTTTATTAACCT 1605
Db      1255 TCTGCTCTGAGAAATGTCATGATGTCCTCAAGGCTTTCGATATGATTTTATTAACCT 1314
Qy      1606 GAGAGTTGAGATTTTCAACCATGAGAGATTAATGTCAGATGATGATTTTGA 1660
Db      1315 GAGAGTTGAGATTTTCAACCATGAGAGATTAATGTCAGATGATGATTTTGA 1369

RESULT 4
AAB33751
ID      AAB33751 standard; DNA; 1400 BP.
XX
AC      AAB33751;
XX
DT      01-JUL-2002 (first entry)
XX
DE      Human MrgX3 (mas-related gene) DNA.
XX
KW      Human; mas-related gene; G-protein coupled receptor; drg-12 protein;
KM      receptor; sensory perception; pain; analgesic; MrgX3; gene; ds.
XX
OS      Homo sapiens.
XX
FH      Key      Location/Qualifiers
FT      CDS      332..1300
FT      /tag= a
FT      /product= "Human MrgX3 protein"
XX
PN      WO200183555-A2.
XX
PD      08-NOV-2001.
XX
PF      04-MAY-2001; 2001WO-US014519.
```

```
XX      04-MAY-2000; 2000US-0202027P.
PR      01-AUG-2000; 2000US-0222344P.
PR      03-NOV-2000; 2000US-00704707.
PR      19-APR-2001; 2001US-0285493P.
XX
PA      (CALY ) CALIFORNIA INST OF TECHNOLOGY.
PI      Anderson DJ, Dong X, Zylka M, Han S, Simon M;
XX      WPI; 2002-171346/22.
DR      P-PSDB; AAE21296.
XX
PT      Isolated polypeptide, Mrg, which is a G-protein coupled receptor and an
PT      isolated polypeptide, drg-12, which is also a receptor, useful for
PT      identifying agonists or antagonists for treating pain.
XX
PS      Disclosure; Page 128-129; 185pp; English.
XX
CC      The invention relates to Mrg (mas-related gene) protein, which is a G-
CC      protein coupled receptor and drg-12 protein, which is a receptor. The
CC      invention is useful for identifying compounds that bind to it, especially
CC      agonists or antagonists. Administration of an agent (e.g. the identified
CC      agonist) that increases the expression of Mrg in a mammal may be used for
CC      treating impaired sensory perception in a mammal, especially pain. The
CC      antagonist may also be useful for treating impaired sensory perception in
CC      a mammal. The present sequence is human MrgX3 DNA
XX
SQ      Sequence 1400 BP; 276 A; 375 C; 363 G; 386 T; 0 U; 0 Other;
XX
Query Match      41.6%; Score 1088; DB 6; Length 1400;
Best Local Similarity 100.0%; Pred. No. 0;
Matches 1088; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy      473 CAGATGGGGTTTCTGAGCATGATTAACCATCCAGTTGGGTACAGAACTGACACC 532
Db      313 CAGATGGGGTTTCTGAGCATGATTAACCATCCAGTTGGGTACAGAACTGACACC 372
Qy      533 AATCAACGAGCGTGAAGAGATCTCTGCTCAAGAGAGACCTGAGCTTCAAGGGGCTGAC 592
Db      373 AATCAACGAGCGTGAAGAGATCTCTGCTCAAGAGAGACCTGAGCTTCAAGGGGCTGAC 432
Qy      593 GTGCATGTTTCCCTTGTCCGCTGACAGAGAAAGCGGTTGTGCTGCTGCTGAGCTG 652
Db      433 GTGCATGTTTCCCTTGTCCGCTGACAGAGAAAGCGGTTGTGCTGCTGAGCTG 492
Qy      493 CCGATGCGCAGAGAGCGTGTCTCAATCTCAATCTCAACCTGATGCGCGGCACTTCT 552
Db      653 CCGATGCGCAGAGAGCGTGTCTCAATCTCAATCTCAACCTGATGCGCGGCACTTCT 712
Qy      713 CTTCCTTAGCGGCACTTAATATGTTGCGGTTACAGCTCATTAATATGCGCCATCCCAT 772
Db      553 CTTCCTTAGCGGCACTTAATATGTTGCGGTTACAGCTCATTAATATGCGCCATCCCAT 612
Qy      773 CTCGAAATCTCAGTCTGATGATGACCTTCCCACTTAATATGAGCTTAAGATGCTGAG 832
Db      613 CTCGAAATCTCAGTCTGATGATGACCTTCCCACTTAATATGAGCTTAAGATGCTGAG 672
Qy      833 CGGCATGACAGACGAGGCTGCTGTCCATCTGTGGCCCATGTGTACCATGCGCGG 892
Db      673 CGGCATGACAGACGAGGCTGCTGTCCATCTGTGGCCCATGTGTACCATGCGCGG 732
Qy      893 CCCGATACCTGTCAATGATGATGTCCTGTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 952
Db      733 CCCGATACCTGTCAATGATGATGTCCTGTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 792
Qy      953 TATCTGGAGTGAATGTTCTGATCTTCCGTTTAAAGTGAAGTGAATCTGTTGGTGTGA 1012
Db      793 TATCTGGAGTGAATGTTCTGATCTTCCGTTTAAAGTGAAGTGAATCTGTTGGTGTGA 852
Qy      1013 AACGTGATTTTATTAACAATTCGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 1072
Db      853 AACGTGATTTTATTAACAATTCGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 912
```

QY 1073 CAGCTGTGCTCTGTGTGAGATTCCTGTGATCCCGAAGATCCGCTGACAGGCT 1132
 DB 913 CAGCTGTGCTCTGTGTGAGATTCCTGTGATCCCGAAGATCCGCTGACAGGCT 972
 QY 1133 GTAGTGAACATCCCTCCACAGATGCTGCTCTCCCTCTGTGCTGCTGCTGCTGCT 1192
 DB 973 GTAGTGAACATCCCTCCACAGATGCTGCTCTCCCTCTGTGCTGCTGCTGCTGCT 1032
 QY 1193 TCAGTGGGCTCTGTGTGAGATTCCTGTGATCCCGAAGATCCGCTGACAGGCT 1252
 DB 1033 TCAGTGGGCTCTGTGTGAGATTCCTGTGATCCCGAAGATCCGCTGACAGGCT 1092
 QY 1253 TCAGTGGGCTCTGTGTGAGATTCCTGTGATCCCGAAGATCCGCTGACAGGCT 1312
 DB 1093 TCAGTGGGCTCTGTGTGAGATTCCTGTGATCCCGAAGATCCGCTGACAGGCT 1152
 QY 1313 CAGTGGGCTCTGTGTGAGATTCCTGTGATCCCGAAGATCCGCTGACAGGCT 1372
 DB 1153 CAGTGGGCTCTGTGTGAGATTCCTGTGATCCCGAAGATCCGCTGACAGGCT 1212
 QY 1373 TCAGTGGGCTCTGTGTGAGATTCCTGTGATCCCGAAGATCCGCTGACAGGCT 1432
 DB 1213 TCAGTGGGCTCTGTGTGAGATTCCTGTGATCCCGAAGATCCGCTGACAGGCT 1272
 QY 1433 CAGTGGGCTCTGTGTGAGATTCCTGTGATCCCGAAGATCCGCTGACAGGCT 1492
 DB 1273 CAGTGGGCTCTGTGTGAGATTCCTGTGATCCCGAAGATCCGCTGACAGGCT 1332
 QY 1493 CAGTGGGCTCTGTGTGAGATTCCTGTGATCCCGAAGATCCGCTGACAGGCT 1552
 DB 1333 CAGTGGGCTCTGTGTGAGATTCCTGTGATCCCGAAGATCCGCTGACAGGCT 1392
 QY 1553 CAGTGGGCTCTGTGTGAGATTCCTGTGATCCCGAAGATCCGCTGACAGGCT 1512
 DB 1393 CAGTGGGCTCTGTGTGAGATTCCTGTGATCCCGAAGATCCGCTGACAGGCT 1400

RESULT 5

AB242595
 ID AB242595 standard; DNA; 1400 BP.

AC AB242595;

DT 04-MAR-2003 (first entry)

DE Human G protein-coupled receptor MrgX3 nucleotide SEQ ID NO:673.

KM G protein-coupled receptor; GPCR; antigenic peptide; gene therapy;
 KM G protein-coupled receptor; modulator; antibody; immune-related disease;
 KM growth-related disease; cell regeneration-related disease; AIDS; cancer;
 KM immunological-related cell proliferative disease; autoimmune disease;
 KM Alzheimer's disease; atherosclerosis; infection; osteoarthritis; allergy;
 KM osteoporosis; cardiomyopathy; inflammation; Crohn's disease; diabetes;
 KM graft versus host disease; Parkinson's disease; multiple sclerosis; pain;
 KM psoriasis; anxiety; depression; schizophrenia; dementia; memory loss;
 KM mental retardation; epilepsy; asthma; tuberculosis; obesity; nausea;
 KM hypertension; hypotension; renal disorder; rheumatoid arthritis; trauma;
 KM ulcer; gene; ds.

OS Homo sapiens.

PN WO200261087-A2.

PD 08-AUG-2002.

PF 19-DEC-2001; 2001WO-US050107.

PR 19-DEC-2000; 2000US-0257144P.

XX (LIFE-) LIFESPAN BIOSCIENCES INC.

XX Burner GC, Roush CL, Brown JP;

XX WPI; 2003-046718/04.
 DR P-PSDB; ABP81750.
 XX New isolated antigenic peptides e.g., for G protein-coupled receptors
 PT (GPCR), useful for diagnosing and designing drugs for treating conditions
 PT in which GPCR are involved, e.g. AIDS, Alzheimer's disease, cancer or
 PT autoimmune diseases.
 XX Disclosure; Fig 1; 523bp; English.
 XX
 CC The present invention describes antigenic peptides (i) comprising: (a)
 CC any one of 1601 sequences (see ABP82019 to ABP83619) of 12-24 amino
 CC acids. Also described: (1) an assay for the detection of a particular G
 CC protein-coupled receptor (GPCR) or a candidate polypeptide in a sample;
 CC and (2) an isolated antibody having high specificity and high affinity or
 CC avidity for a particular GPCR. (i) can be used as GPCR modulators and in
 CC gene therapy. The antigenic peptides for GPCR are useful in detecting an
 CC antibody against a particular GPCR, and in the production of specific
 CC antibodies. The peptides and antibodies are also useful for detecting the
 CC presence or absence of corresponding GPCRs. The antigenic peptides for
 CC GPCRs and antibodies are useful for diagnosing and designing drugs for
 CC treating immune-related diseases, growth-related diseases, cell
 CC regeneration-related diseases, immunological-related cell proliferative
 CC diseases, or autoimmune diseases, e.g. AIDS, Alzheimer's disease,
 CC atherosclerosis, bacterial, fungal, protozoan or viral infections,
 CC osteoarthritis, osteoporosis, cancer, cardiomyopathy, chronic and acute
 CC inflammation, allergies, Crohn's disease, diabetes, graft versus host
 CC disease, Parkinson's disease, multiple sclerosis, pain, psoriasis,
 CC anxiety, depression, schizophrenia, dementia, mental retardation, memory
 CC loss, epilepsy, asthma, tuberculosis, obesity, nausea, hypertension,
 CC hypotension, renal disorders, rheumatoid arthritis, trauma, ulcer, or
 CC any other disorder in which GPCRs are involved. The antibodies may be
 CC used in immunoassays and immunodiagnosis. AB242523 to AB242869 encode
 CC GPCR proteins given in ABP81675 to ABP82018, which are used in the
 CC exemplification of the present invention
 XX
 SQ Sequence 1400 BP; 276 A; 375 C; 363 G; 386 T; 0 U; 0 Other;

Query Match 41.6%; Score 1088; DB 8; Length 1400;
 Best Local Similarity 100.0%; Pred. No. 0;
 Matches 1088; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 473 CAGACTGGGGTTTCTGACATGATTCACCATCCAGTCTTGGGTACAGAACTGACACC 532
 DB 313 CAGACTGGGGTTTCTGACATGATTCACCATCCAGTCTTGGGTACAGAACTGACACC 372
 QY 533 AATCAACGAGCGTGAAGAGATCTCTGTGTAACAGCAACCTGAGCTTACAGGGGCTGAC 592
 DB 373 AATCAACGAGCGTGAAGAGATCTCTGTGTAACAGCAACCTGAGCTTACAGGGGCTGAC 432
 QY 593 GTGATCGTTTCCCTTGGCGGCTGACAGAAACGGGTTGTGCTCTGAGCTCTGGGCTG 652
 DB 433 GTGATCGTTTCCCTTGGCGGCTGACAGAAACGGGTTGTGCTCTGAGCTCTGGGCTG 492
 QY 653 CCGCATGCGAGAAAGCTGTCTCCATCTACATCTCACTCACTGAGCGGCGACTTCT 712
 DB 493 CCGCATGCGAGAAAGCTGTCTCCATCTACATCTCACTCACTGAGCGGCGACTTCT 552
 QY 713 CTTCCTTACGGGCAATTAATGTTGGCGGTTAGCGCTCAATCAATTCGGCATTCAT 772
 DB 553 CTTCCTTACGGGCAATTAATGTTGGCGGTTAGCGCTCAATCAATTCGGCATTCAT 612
 QY 773 CTTCCTTACGGGCAATTAATGTTGGCGGTTAGCGCTCAATTAATTCGGCATTCAT 832
 DB 613 CTTCCTTACGGGCAATTAATGTTGGCGGTTAGCGCTCAATTAATTCGGCATTCAT 672
 QY 833 CGCCATGAGCAACGAGCGCTGCTGCTCACTCTGAGCGCATCTGTAGCACTGCGCG 892
 DB 673 CGCCATGAGCAACGAGCGCTGCTGCTCACTCTGAGCGCATCTGTAGCACTGCGCG 732
 QY 893 CCCGATACCTGTGATGCTGATGTTGTTCTGCTGAGCGCTGCTGCTGCTGCTGCTG 952

QY 1133 GTACGATGACATCTCTCTCAAGTGTGTCTTCTCTCTGTGTGCTCCCTTTGGCAT 1192
DB 973 GTACGATGACATCTCTCTCAAGTGTGTCTTCTCTCTGTGTGCTCCCTTTGGCAT 1032
QY 1193 TCAGTGGGCGCTTTTTCAGATTCACCTGATTTGGAAGTCTATTTTTCAGATGCA 1252
DB 1033 TCAGTGGGCGCTTTTTCAGATTCACCTGATTTGGAAGTCTATTTTTCAGATGCA 1092
QY 1253 TCTAGTTTCCATTTTCTGTGCTCTTAAACAGAGTGCACCCCATATTACTTCTT 1312
DB 1093 TCTAGTTTCCATTTTCTGTGCTCTTAAACAGAGTGCACCCCATATTACTTCTT 1152
QY 1313 CTGTGGGCTCTTTTTCAGAGTGCACCTGAAATGCAAACTGAACTGTCTTCCAGAGGCG 1372
DB 1153 CTGTGGGCTCTTTTTCAGAGTGCACCTGAAATGCAAACTGAACTGTCTTCCAGAGGCG 1212
QY 1373 TCTGACAGACACGCTGAGTGTGATGAAAGTGAAGGCTTCCCTGAGAAACCTGTGA 1432
DB 1213 TCTGACAGACACGCTGAGTGTGATGAAAGTGAAGGCTTCCCTGAGAAACCTGTGA 1272
QY 1433 GCTGTGCGGAAGCAGATTTGAGCAGTGAAGAACTTGTGCTGTCTGACAGAGACTT 1492
DB 1273 GCTGTGCGGAAGCAGATTTGAGCAGTGAAGAACTTGTGCTGTCTGACAGAGACTT 1332
QY 1493 GAGAGCAATGCTGCTGCTGACCTGACCTTGAATTAATGCAATTTTTCAGCTTGTGCT 1552
DB 1333 GAGAGCAATGCTGCTGCTGACCTGACCTTGAATTAATGCAATTTTTCAGCTTGTGCT 1392
QY 1553 CAGAAATG 1560
DB 1393 CAGAAATG 1400

RESULT 7
AAA29811

ID AAA29811 standard; cDNA; 969 BP.

AC AAA29811;

DT 18-AUG-2000 (first entry)

DE Human G protein-coupled receptor hH17213 encoding cDNA SEQ ID NO:3.

KW Human; G protein-coupled receptor; hippocampus; diagnosis; screening;
genetic disease; cellular function regulation; ss.

OS Homo sapiens.

FH Key Location/Qualifiers

FT CDS 1..969

FT /tag= a

FT /product= "G protein-coupled receptor"

PN MO200020455-A1.

XX 13-APR-2000.

XX 30-SEP-1999; 99MO-JP005366.

XX 01-OCT-1998; 98JP-00279535.

PA (TAKE) TAKEDA CHEM IND LTD.

PI Watanabe T, Terao Y, Matsu H;

DR WPI: 2000-303747/26.

XX P-PSDB: AA190761.

PT Human-derived G protein-coupled protein and encoding nucleic acid, useful
in determining ligands and treatment of diseases associated with
dysfunction of the protein.

PS Claim 6; Page 93-94; 97pp; Japanese.

XX The present sequence encodes a human-derived G protein-coupled protein
CC designated hH17213, which is isolated from the human hippocampus. The G
CC protein-coupled receptor can be used for preventing, treating and
CC diagnosing genetic diseases associated with G protein-coupled protein.
CC and for regulating cellular functions. The protein can be used to prevent
CC and treat disorders associated with G protein-coupled protein gene
CC dysfunction. It can also be used to identify G protein-coupled protein
CC ligands and generating antibodies and antisera against the protein. It is
CC also useful in constructing recombinant receptor protein expression
CC systems, developing receptor-binding assay systems and screening drug
CC candidates, and can be used as a probe in the genetic diagnosis of G
CC protein-coupled protein disorders

SQ Sequence 969 BP; 171 A; 280 C; 250 G; 268 T; 0 U; 0 Other;

Query Match 37.0%; Score 969; DB 3; Length 969;
Best Local Similarity 100.0%; Pred. No. 0;
Matches 969; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 492 ATGATTCACCAATCCAGTCTTGGGTAAGAACTGACCAATCAACGAGAGAG 551
DB 1 ATGATTCACCAATCCAGTCTTGGGTAAGAACTGACCAATCAACGAGAGAG 60
QY 552 ACTCTGTGTAACAGACACCTGAGCTTCAAGGAGGCTGAGTGCATGCTTCCCTGTC 611
DB 61 ACTCTGTGTAACAGACACCTGAGCTTCAAGGAGGCTGAGTGCATGCTTCCCTGTC 120
QY 612 GCGCTGACAGAAAGCGGTTGTGCTTGTGCTCTGAGCTGAGGCTGAGGAGAGAGCT 671
DB 121 GCGCTGACAGAAAGCGGTTGTGCTTGTGCTCTGAGCTGAGGCTGAGGAGAGAGCT 180
QY 672 GTTCATCATCATCTCAACCTGTGCGGCGCACTTCTCTTCTTACGCGGCACTT 731
DB 181 GTTCATCATCATCTCAACCTGTGCGGCGCACTTCTCTTCTTACGCGGCACTT 240
QY 732 ATATGTTGCGCGCTTACGCTCATCAATATCCGCCATCCCATCTCCAAATCTCAATGCT 791
DB 241 ATATGTTGCGCGCTTACGCTCATCAATATCCGCCATCCCATCTCCAAATCTCAATGCT 300
QY 792 GTGATGACCTTCCCTCACTTATAGGCTTAAGAGCATGAGGAGGAGGAGGAGGAG 851
DB 301 GTGATGACCTTCCCTCACTTATAGGCTTAAGAGCATGAGGAGGAGGAGGAGGAGGAG 360
QY 852 TGCCTGTCCATCTGTGCGGCACTGTGATACACCTGCGGCGGCGGAGATCTGTGATG 911
DB 361 TGCCTGTCCATCTGTGCGGCACTGTGATACACCTGCGGCGGCGGAGATCTGTGATG 420
QY 912 GTGATGATGCTGCTGCTGCGGCGGCTGTGCTGCTGCGGAGATATCTGAGATGATGTC 971
DB 421 GTGATGATGCTGCTGCTGCGGCGGCTGTGCTGCTGCGGAGATATCTGAGATGATGTC 480
QY 972 TGTGACTTCTTGTAGTGTGCTGATCTGTGTTGTGGAACGTCAGATTTCATTA 1031
DB 481 TGTGACTTCTTGTAGTGTGCTGATCTGTGTTGTGGAACGTCAGATTTCATTA 540
QY 1032 ATGCGGTGCTGTTTATGATGATGATGATGATGATGATGATGATGATGATGATGATG 1091
DB 541 ATGCGGTGCTGTTTATGATGATGATGATGATGATGATGATGATGATGATGATGATG 600
QY 1092 AGGATTTCTGTGATATCCCGAAGATGCGGCTGACACAGGCTGTAGCTGATCACTCTCTC 1151
DB 601 AGGATTTCTGTGATATCCCGAAGATGCGGCTGACACAGGCTGTAGCTGATCACTCTCTC 660
QY 1152 ACAGTGTGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATG 1211
DB 661 ACAGTGTGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATG 720
QY 1212 AGGATTCACCTGATGAAAGTCTTATTTGTGATGATGATGATGATGATGATGATGATG 1271
DB 721 AGGATTCACCTGATGAAAGTCTTATTTGTGATGATGATGATGATGATGATGATGATG 780
QY 1272 TCCGCTTTACAGAGAGTGCACCAACCATCATTTACTTCTGTGAGGCTCTTTAGGAG 1331

```
Db 781 TCGGCTTTAAACAGACAGTGCACCAACCCATCATTTACTTCTTCGGGGCTCCTTTAGGAG 840
Qy 1332 CGTCAAAATATGGAGAACTGGAAGCTGTTCTCCAGAGGGCTCTGCAGAGACAGCCCTGAG 1391
Db 841 CGTCAAAATATGGAGAACTGGAAGCTGTTCTCCAGAGGGCTCTGCAGAGACAGCCCTGAG 900
Qy 1392 GTGGAATGAAGTGAAGGGTGGCTTCTCTCAGAGAAACCTTGAGCTGTGCGAGACAGATTG 1451
Db 901 GTGGAATGAAGTGAAGGGTGGCTTCTCTCAGAGAAACCTTGAGCTGTGCGAGACAGATTG 960
Qy 1452 GAGCAGTGA 1460
Db 961 GAGCAGTGA 969

RESULT 8
ABT04875
ID ABT04875 standard; cDNA; 969 BP.
XX
AC ABT04875;
XX
DT 11-OCT-2002 (first entry)
XX
DE Human G protein coupled receptor hRUP37 coding sequence.
XX
KW Human; G-protein coupled receptor; GPCR; hRUP28; hRUP29; hRUP30; hRUP31;
KW hRUP32; hRUP33; hRUP34; hRUP35; hRUP36; hRUP37; gene; ss.
XX
OS Homo sapiens.
XX
PN M0200242461-A2.
XX
PD 30-MAY-2002.
XX
PF 26-NOV-2001; 2001MO-US044386.
XX
PR 27-NOV-2000; 2000US-0253404P.
PR 12-DEC-2000; 2000US-0255366P.
PR 20-FEB-2001; 2001US-0270286P.
PR 20-FEB-2001; 2001US-0270286P.
PR 06-APR-2001; 2001US-0282032P.
PR 06-APR-2001; 2001US-0282356P.
PR 06-APR-2001; 2001US-0282358P.
PR 06-APR-2001; 2001US-0282365P.
PR 14-MAY-2001; 2001US-0290917P.
PR 31-JUL-2001; 2001US-0309208P.
XX
PA (AREN-) ARENA PHARM INC.
XX
PI Chen R, Chu ZL, Dang HT, Lowitz KP, Pride C;
XX
DR MPI; 2002-566565/60.
XX
PT P-PSDB; ABJ04077.
XX
PT Novel endogenous and non-endogenous versions of G protein-coupled
PT receptor useful for identification of candidate compounds as receptor
PT agonists or antagonists for use as therapeutic agents.
XX
PS Claim 39; Page 74; 84pp; English.
XX
CC The present invention provides the protein and coding sequences of
CC several human G-protein coupled receptors (GPCRs). These can be used in
CC the identification of candidate compounds as receptor agonists or inverse
CC agonists having applicability as therapeutic agents. The present sequence
CC is a GPCR coding sequence of the invention.
XX
SO Sequence 969 BP; 171 A; 280 C; 250 G; 268 T; 0 U; 0 Other;
```

```
Query Match 37.0%; Score 969; DB 6; Length 969;
Best Local Similarity 100.0%; Pred. No. 0;
Matches 969; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
```

```
Qy 492 ATGATTCACCATCCAGTCTTGGGTACAGAACTGACCAATTCACGAGCTGAGAG 551
Db 1 ATGATTCACCATCCAGTCTTGGGTACAGAACTGACCAATTCACGAGCTGAGAG 60
Qy 552 ACTCTTGTCTACAGACAGACCTTGAAGTTTCAAGGGGCTGACGTGTCATCTTCCCTTGTG 611
Db 61 ACTCTTGTCTACAGACAGACCTTGAAGTTTCAAGGGGCTGACGTGTCATCTTCCCTTGTG 120
Qy 612 GCGGTGACAGAAACGGGTTGTGCTTGGCTCTGGGCTCGGGCTGCGCATCGCAGAGAGCT 671
Db 121 GCGGTGACAGAAACGGGTTGTGCTTGGCTCTGGGCTCGGGCTGCGCATCGCAGAGAGCT 180
Qy 672 GTCTCCATTCATACCTTCAACCTGATCGCGCCGACCTTCTCTTCTTACCGGACATT 731
Db 181 GTCTCCATTCATACCTTCAACCTGATCGCGCCGACCTTCTCTTCTTACCGGACATT 240
Qy 732 ATATGTTGCGGTTAGGCGCTCATCATATATCCGCATTCCTTCCAAATCTTCAGTCT 791
Db 241 ATATGTTGCGGTTAGGCGCTCATCATATATCCGCATTCCTTCCAAATCTTCAGTCT 300
Qy 792 GTGATGACCTTTCCTACTTATATAGGCTTAAGCATGCTGAGCGCATCAGACCGAGCGG 851
Db 301 GTGATGACCTTTCCTACTTATATAGGCTTAAGCATGCTGAGCGCATCAGACCGAGCGG 360
Qy 852 TGCCTGTCCATCCTGTGAGCCCATCTGATACCACTGCGCGCCGACAGTACTGTCAATG 911
Db 361 TGCCTGTCCATCCTGTGAGCCCATCTGATACCACTGCGCGCCGACAGTACTGTCAATG 420
Qy 912 GTCATGTGTGCTGCTCTGAGCCCTGTGCTCTGCTGCGAGTATCTGAGTGTGATTTTC 971
Db 421 GTCATGTGTGCTGCTCTGAGCCCTGTGCTCTGCTGCGAGTATCTGAGTGTGATTTTC 480
Qy 972 TGTGACTTCTGTTTATGTTGTTGTTGTTGTTGTTGTTGTTGTTGTTGTTGTTGTTGTTG 1031
Db 481 TGTGACTTCTGTTTATGTTGTTGTTGTTGTTGTTGTTGTTGTTGTTGTTGTTGTTGTTG 540
Qy 1032 ATGCGGTGCTGTTTATGTTGTTGTTGTTGTTGTTGTTGTTGTTGTTGTTGTTGTTGTTG 1091
Db 541 ATGCGGTGCTGTTTATGTTGTTGTTGTTGTTGTTGTTGTTGTTGTTGTTGTTGTTGTTG 600
Qy 1092 AGGATTTCTGTGATATCCGGAAGATGCGCTGACCAAGCTGTATCTGATCAATCTTCTTC 1151
Db 601 AGGATTTCTGTGATATCCGGAAGATGCGCTGACCAAGCTGTATCTGATCAATCTTCTTC 660
Qy 1152 ACAGTGTGTTCTTCTCTCTGTCGTCGTCGTCGTCGTCGTCGTCGTCGTCGTCGTCGTCG 1211
Db 661 ACAGTGTGTTCTTCTCTCTGTCGTCGTCGTCGTCGTCGTCGTCGTCGTCGTCGTCGTCG 720
Qy 1212 AGGATCCACTGATTTGAAAGTCTTATTTTGTATGATGTCATCTAGTTTCCATTTTCTTG 1271
Db 721 AGGATCCACTGATTTGAAAGTCTTATTTTGTATGATGTCATCTAGTTTCCATTTTCTTG 780
Qy 1272 TCGGCTTTTAAACAGACGTGCAACCCCATCATTTACTTCTTGTGGGCTCCTTTAGGAG 1331
Db 781 TCGGCTTTTAAACAGACGTGCAACCCCATCATTTACTTCTTGTGGGCTCCTTTAGGAG 840
Qy 1332 CGTCAAAATATGGAGAACTGGAAGCTGTTCTCCAGAGGGCTCTGCAGAGACAGCCCTGAG 1391
Db 841 CGTCAAAATATGGAGAACTGGAAGCTGTTCTCCAGAGGGCTCTGCAGAGACAGCCCTGAG 900
Qy 1392 GTGGAATGAAGTGAAGGGTGGCTTCTCTCAGAGAAACCTTGAGCTGTGCGAGACAGATTG 1451
Db 901 GTGGAATGAAGTGAAGGGTGGCTTCTCTCAGAGAAACCTTGAGCTGTGCGAGACAGATTG 960
Qy 1452 GAGCAGTGA 1460
Db 961 GAGCAGTGA 969
```

```
RESULT 9
ADO30080
ID ADO30080 standard; cDNA; 969 BP.
XX
```


AC ADO30080;
 XX 29-JUL-2004 (first entry)
 XX
 DE Human GPCR MEGX3 polynucleotide, SEQ ID NO:1182.
 XX
 KM G protein-coupled receptor; GPCR; drug screening; diagnosis;
 KM transgenic mouse; neurological disorder; adrenal gland disorder;
 KM colon disorder; intestinal disorder; cardiovascular disorder;
 KM muscular disorder; blood disorder; immune disorder; bone disorder;
 KM joint disorder; metabolic disorder; nutritive disorder; cancer;
 KM kidney disorder; liver disorder; lung disorder; breast disorder;
 KM ovary disorder; uterus disorder; prostate disorder; testis disorder;
 KM skin disorder; stomach disorder; pancreas disorder; spleen disorder;
 KM thymus disorder; thyroid disorder; antiparkinsonian; autonomic;
 KM cytoskeletal; antiinflammatory; vasotropic; antidiarrhoeic;
 KM CNS; central nervous system; respiratory; antidiarrhoeic; antidiabetic;
 KM virucide; hepatotropic; antibacterial; antianemic; antiseborrhoeic;
 KM dermatological; antitumor; antihypertensive; anorectic;
 KM immunosuppressive; nephrotropic; gene therapy; GPCR modulator; human;
 KM gene; ss.
 XX
 OS Homo sapiens.
 XX
 PN WO200404000-A2.
 XX
 PD 13-MAY-2004.
 XX
 PF 09-SEP-2003; 2003WO-US028226.
 XX
 PR 09-SEP-2002; 2002US-0409303P.
 PR 09-APR-2003; 2003US-0461329P.
 XX
 XX (PRIM-) PRIMAL INC.
 XX Gaftanaris GA, Bergmann JE, Gragerov A, Hohmann J, Li F,
 PI Madisen L, McIlwain KL, Pavlova MN, Vassiliadis D, Zeng H;
 XX
 DR WPI; 2004-390329/36.
 XX P-PSDB; ADO29705.
 XX
 PT Novel mammalian G protein coupled receptors, useful for identifying
 PT compounds that modulates diagnosing and treating disease condition
 PT associated with GPCR dysfunction e.g. autoimmune diseases, angina
 PT pectoris, Parkinson's disease.
 XX
 PS Claim 151, SEQ ID NO 1182; 542bp; English.
 XX
 CC The invention relates to human and mouse G protein-coupled receptors
 CC (GPCRs) and nucleic acids encoding them. The invention also relates to
 CC sequences at least 90% identical to the GPCR proteins and nucleic acids
 CC of the invention; methods of treating, preventing or diagnosing diseases
 CC associated with GPCRs of the invention; methods of screening for
 CC compounds useful in the treatment of GPCR-related diseases; a transgenic
 CC mouse comprising a GPCR gene or in an endogenous GPCR gene; cells derived
 CC mutation in a GPCR transgene or in an endogenous GPCR gene; cells derived
 CC from the transgenic mice; kits comprising several mice, each of which has
 CC a mutation in a different GPCR gene of the invention; and kits comprising
 CC probes which hybridise to GPCR polynucleotides of the invention. The
 CC invention further discloses variants of the GPCR polypeptides and vectors
 CC comprising a GPCR nucleic acid. The GPCR nucleic acids and proteins may
 CC be used in the diagnosis, treatment or prevention of a wide variety of
 CC diseases including neurological disorders (e.g., Alzheimer's disease,
 CC depression, diabetic neuropathy, Parkinson's disease or schizophrenia);
 CC disorders of the adrenal gland; disorders of the colon or intestine
 CC (e.g., Crohn's disease, diarrhoea, food poisoning or irritable bowel
 CC syndrome); cardiovascular disorders (e.g., angina, cardiac arrhythmia or
 CC myocardial infarction); muscular disorders; blood disorders (e.g.,
 CC anaemia or leukaemia); immune disorders (e.g., autoimmune disorders or
 CC AIDS); bone and joint disorders (e.g., osteoarthritis, rheumatoid
 CC arthritis, gout or osteoporosis); metabolic or nutritive disorders (e.g.,
 CC obesity, enzyme deficiency-related diseases or vitamin deficiency-related
 CC diseases); and disorders of the kidney, liver, lung, breast, ovary,

CC uterus, prostate, testis, skin, stomach, pancreas, spleen, thymus and
 CC thyroid (e.g., cancers). The present sequence represents a GPCR-encoding
 CC nucleic acid of the invention. Note: The full sequence data for this
 CC patent did not form part of the printed specification; those sequences
 CC not shown were obtained in electronic format directly from WIP0 at
 CC ftp.wipo.int/pub/published_pcf_sequences.
 XX
 SQ Sequence 969 BP; 171 A; 280 C; 250 G; 268 T; 0 U; 0 Other;
 XX
 Query Match 37.0%; Score 969; DB 12; Length 969;
 Best Local Similarity 100.0%; Pred. No. 0;
 Matches 969; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
 XX
 QY 492 ATGATTCACACCATCCAGTCTTGAGTACAGAACTGACACCAATCAACGAGCTAGAG 551
 1 ATGATTCACACCATCCAGTCTTGAGTACAGAACTGACACCAATCAACGAGCTAGAG 60
 DB
 QY 552 ACTCTTGCTACACGACACCTGAGCTTACAGGGGCTGAGAGTGTCTTCCCTTGTG 611
 61 ACTCTTGCTACACGACACCTGAGCTTACAGGGGCTGAGAGTGTCTTCCCTTGTG 120
 DB
 QY 612 GCGTGAACGAGAAACGCGGTTGTGCTGAGCTCTGAGGCTGCGCATGCGAGAGAGCT 671
 121 GCGTGAACGAGAAACGCGGTTGTGCTGAGCTCTGAGGCTGCGCATGCGAGAGAGCT 180
 DB
 QY 672 GTCTCATCTACATCTCAACCTGCTGCGGCGGCACTTCTTCTTTAGCGGCACTT 731
 181 GTCTCATCTACATCTCAACCTGCTGCGGCGGCACTTCTTCTTTAGCGGCACTT 240
 DB
 QY 732 ATATGTTGGCGGTACGCTCATCAATATCCGCCATCCCATCTCCAAATCTCAATCT 791
 241 ATATGTTGGCGGTACGCTCATCAATATCCGCCATCCCATCTCCAAATCTCAATCT 300
 DB
 QY 792 GTGATGACCTTCCCTTATAGGCTTAAGCATGCTGAGCGCATGAGCAGCAGCGC 851
 301 GTGATGACCTTCCCTTATAGGCTTAAGCATGCTGAGCGCATGAGCAGCAGCGC 360
 DB
 QY 852 TGCCTGTCCATCTGTGAGCCCATCTGTACCACTGCGCGGCGCCCAATACCTGTATCG 911
 361 TGCCTGTCCATCTGTGAGCCCATCTGTACCACTGCGCGGCGCCCAATACCTGTATCG 420
 DB
 QY 912 GTCATGTGTGCTGCTGCTGCGGCGGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 971
 421 GTCATGTGTGCTGCTGCTGCGGCGGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 480
 DB
 QY 972 TGTGACTTCCTGTTAGTGTGCTGATCTGTGTTGTTGTTGTTGTTGTTGTTGTTG 1031
 481 TGTGACTTCCTGTTAGTGTGCTGATCTGTGTTGTTGTTGTTGTTGTTGTTGTTGTTG 540
 DB
 QY 1032 ATCGGCTGAGCTGTTTATATGTTGTTGTTGTTGTTGTTGTTGTTGTTGTTGTTG 1091
 541 ATCGGCTGAGCTGTTTATATGTTGTTGTTGTTGTTGTTGTTGTTGTTGTTGTTGTTG 600
 DB
 QY 1092 AGGATTTCTGTGATATCCCGGAAATGCGCTGACAGAGCTGTATGATCAATCTCTCTC 1151
 601 AGGATTTCTGTGATATCCCGGAAATGCGCTGACAGAGCTGTATGATCAATCTCTCTC 660
 DB
 QY 1152 AAGAGTGTGTTCTTCTCTCTGTTGAGCTGCTTGTGATTCAGTGGGCTCTTTTTC 1211
 661 AAGAGTGTGTTCTTCTCTCTGTTGAGCTGCTTGTGATTCAGTGGGCTCTTTTTC 720
 DB
 QY 1212 AGGATTCACCTGATGAGAAAGCTTATTTTGTATGTCATGATGATTTTCTCTG 1271
 721 AGGATTCACCTGATGAGAAAGCTTATTTTGTATGTCATGATGATTTTCTCTG 780
 DB
 QY 1272 TCCGCTTTTAAAGAGAGGCCAACCCCATCAATTAATCTTCTGAGGCTCTTTAGGAG 1331
 781 TCCGCTTTTAAAGAGAGGCCAACCCCATCAATTAATCTTCTGAGGCTCTTTAGGAG 840
 DB
 QY 1332 CGTCAAAATGAGCAGAACTGATGAGCTGTTCTGAGAGGGCTGACAGACCGCTGAG 1391
 841 CGTCAAAATGAGCAGAACTGATGAGCTGTTCTGAGAGGGCTGACAGACCGCTGAG 900
 DB

QY 1392 GTGATGATGAGTGGAGGCTTCTCAGAGAAACCTGAGCTGTGGAGAGAGATTG 1451
DB 901 GTGATGATGAGTGGAGGCTTCTCAGAGAAACCTGAGCTGTGGAGAGAGATTG 960
QY 1452 GAGCAGTGA 1460
DB 961 GAGCAGTGA 969

RESULT 10
ADO44603
ID ADO44603 standard; DNA; 966 BP.
XX ADO44603;
AC
XX
XX 29-JUL-2004 (first entry)
DT
XX
XX Human HIT7213 protein encoding DNA.
DE
XX
XX HIT7213; transgenic; G protein-coupled receptor; GPCR; ophthalmological;
KM cytostatic; nephrotoxic; antiinflammatory; dermatological; analgesic;
KM vulnary; neuroprotective; human; gene; ds.
XX
XX Homo sapiens.
OS
XX
XX Key Location/Qualifiers
FH 1.966
FT CDS /*tag= a
FT /product= "HIT7213"
XX
XX MO2004039972-A1.
XX
XX 13-MAY-2004.
XX
XX 28-OCT-2003; 2003MO-JP013781.
XX
XX 29-OCT-2002; 2002JP-00314141.
XX
XX (TAKE) TAKEDA CHEM IND LTD.
XX
XX Kaisho Y, Watanabe T, Yasuhara Y, Mori I, Taketomi S,
PI WPI; 2004-376191/35.
XX
XX P-PSDB; ADO44602.
XX
XX HIT7213 protein, encoded DNA and transgenic animals for clarifying
PT patholological mechanism, developing therapeutic methods and screening
PT preventives or remedies for related diseases e.g. cataract, cancer, and
PT dermatitis.
XX
XX
XX Disclosure; SEQ ID NO 2; 161pp; Japanese.
XX
XX The invention relates to a non-human mammal that carries a DNA integrated
CC with a foreign HIT7213 or its mutant gene, or a part of it. The non-human
CC animal is particularly a rat. Such gene shows phenotypes of e.g. cataract
CC onset, transient skin rash and proliferation-promoting activity. The
CC foreign HIT7213 gene is a gene that encodes a G protein-coupled receptor
CC (GPCR) protein HIT7213. The protein, its encoded DNA and constructed
CC transgenic animals are useful for clarifying pathological mechanism,
CC developing therapeutic methods and screening preventives or remedies for
CC related diseases e.g. cataract, cancer, and dermatitis. The present
CC sequence represents a DNA encoding a human HIT7213 protein.
XX
XX Sequence 966 BP; 170 A; 280 C; 249 G; 267 T; 0 U; 0 Other;
SQ

Query Match 36.9%; Score 966; DB 12; Length 966;
Best Local Similarity 100.0%; Pred. No. 0;
Matches 966; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 492 ATGATTCACACATCCAGTCTTGGTACAGAACTGACCAATCAACGACGAGAGAG 551
DB 1 ATGATTCACACATCCAGTCTTGGTACAGAACTGACCAATCAACGACGAGAGAG 60

QY 552 ACTCTTGTCTACAGACAGACCTTGAGCTTACAGGAGGCTGACGTGATCGTTCCCTTGTG 611
DB 61 ACTCTTGTCTACAGACAGACCTTGAGCTTACAGGAGGCTGACGTGATCGTTCCCTTGTG 120
QY 612 GCGCTGACAGAGAAACGGGCTTGTGCTTGTGCTCTGTGGCTGCGGCATGCGCAAGAAAGCT 671
DB 121 GCGCTGACAGAGAAACGGGCTTGTGCTTGTGCTCTGTGGCTGCGGCATGCGCAAGAAAGCT 180
QY 672 GTCCTCATCTACACCTCAACCTGATGCGGCGGAGCTTCTCTTCTTCTTACGCGGCAT 731
DB 181 GTCCTCATCTACACCTCAACCTGATGCGGCGGAGCTTCTCTTCTTCTTACGCGGCAT 240
QY 732 ATATGTTGCGGCTTACGCTCATCATATATCGGCATCCATCTCCAAATCTCTAGTCT 791
DB 241 ATATGTTGCGGCTTACGCTCATCATATATCGGCATCCATCTCCAAATCTCTAGTCT 300
QY 792 GTGATGACCTTTCCTTACTTATAGGCTTATAGCATGTGAGCGGCATCAGACCGAGCGC 851
DB 301 GTGATGACCTTTCCTTACTTATAGGCTTATAGCATGTGAGCGGCATCAGACCGAGCGC 360
QY 852 TGCCTGTCATCCGTTGCGCCCATCTGCTACACGCGCGCGCCAGATACCTGTGATG 911
DB 361 TGCCTGTCATCCGTTGCGCCCATCTGCTACACGCGCGCCAGATACCTGTGATG 420
QY 912 GTCATGTGTCTCTGCTCTGAGGCTCTGCTCCCTGCGAGATATCTGAGTGTGATTC 971
DB 421 GTCATGTGTCTCTGCTCTGAGGCTCTGCTCCCTGCGAGATATCTGAGTGTGATTC 480
QY 972 TGTGACTTCTCTGTTATGAGTGTGCTGATCTGTTTGTGTGAACGTGATTTTATTA 1031
DB 481 TGTGACTTCTCTGTTATGAGTGTGCTGATCTGTTTGTGTGAACGTGATTTTATTA 540
QY 1032 ATGCGGTGCTGATTTTATGATGATGCTCTGAGGCTGCAAGCCGTGCTGCTGATG 1091
DB 541 ATGCGGTGCTGATTTTATGATGATGCTCTGAGGCTGCAAGCCGTGCTGCTGATG 600
QY 1092 AGGATTCCTGTGAGATCCCGAAGATGCGCTGACAGAGCTGTATCGTATCATCTCTC 1151
DB 601 AGGATTCCTGTGAGATCCCGAAGATGCGCTGACAGAGCTGTATCGTATCATCTCTC 660
QY 1152 ACAAGTCTGTCTCTCTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 1211
DB 661 ACAAGTCTGTCTCTCTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 720
QY 1212 AGGATTCACCTGATTTGGAAGTCTTATTTGTCATGTCGATCATGTTCCATTTTCTG 1271
DB 721 AGGATTCACCTGATTTGGAAGTCTTATTTGTCATGTCGATCATGTTTCCATTTTCTG 780
QY 1272 TCCGCTCTTAACAGACAGTGCACCCCATCATTTACTTCTGTGAGCTCTTTAGGAG 1331
DB 781 TCCGCTCTTAACAGACAGTGCACCCCATCATTTACTTCTGTGAGCTCTTTAGGAG 840
QY 1332 CGTCAAAATAGGAGAACTGAAAGCTGTTCTTCAGAGGCTCTGACAGACGCTGAG 1391
DB 841 CGTCAAAATAGGAGAACTGAAAGCTGTTCTTCAGAGGCTCTGACAGACGCTGAG 900
QY 1392 GTGATGATGAGTGGAGGCTTCTCAGAGAAACCTGAGCTGTGGAGAGAGATTG 1451
DB 901 GTGATGATGAGTGGAGGCTTCTCAGAGAAACCTGAGCTGTGGAGAGAGATTG 960
QY 1452 GAGCAG 1457
DB 961 GAGCAG 966

RESULT 11
ADP70583
ID ADP70583 standard; DNA; 1683 BP.
XX
XX ADP70583;
XX
XX 12-FEB-2004 (first entry)
XX

Orphan receptor ligand-related human protein gene SeqID206.

XX ligand: orphan receptor protein; fusion protein; fluorescent protein;
 XX cell expression; green fluorescent protein; GFP; GFP-1; wild-type GFP;
 KM GFPuv; Enhanced GFP; EGFP; human; gene; ds.
 XX
 OS Homo sapiens.
 PN MO2003071272-A1.
 XX
 PD 28-AUG-2003.
 XX
 PF 21-FEB-2003; 2003WO-JP001901.
 XX
 PR 22-FEB-2002; 2002JP-00045728.
 PR 23-JUL-2002; 2002JP-00213949.
 PR 11-OCT-2002; 2002JP-00298237.
 XX
 PA (TAKE) TAKEDA CHEM IND LTD.
 XX
 PI Hinuma S, Fujii R, Ogi K, Komatsu H, Kawamata Y, Hosoya M;
 XX
 DR WPI; 2003-697654/66.
 DR P-PSDB; ADF70481.
 XX
 PT Transformation of cells with a fusion protein of an orphan receptor
 PT protein with a fluorescent protein useful for identification of ligands
 PT to the orphan receptor.
 XX
 PS Example 4; SEQ ID NO 206; 594bp; Japanese.
 XX
 CC This invention relates to a novel method of identifying ligands to an
 CC orphan receptor protein which comprises transforming cells with DNA
 CC encoding a fusion protein of the orphan receptor with a fluorescent
 CC protein, so that the fusion protein is expressed in the cells (or cell
 CC membranes isolated from them) and contacting the cells with the potential
 CC ligand to be tested. A suitable fluorescent protein for incorporation in
 CC the fusion protein is green fluorescent protein (GFP), for example GFP-1,
 CC wild-type GFP, GFPuv or Enhanced GFP (EGFP). The method is useful for the
 CC identification of ligands binding to an orphan receptor protein.
 CC
 SQ Sequence 1683 BP; 402 A; 438 C; 401 G; 442 T; 0 U; 0 Other;

Query Match 36.9%; Score 966; DB 10; Length 1683;
 Best Local Similarity 100.0%; Pred. No. 0;
 Matches 966; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 492 ATGATTCACCATTCCTGCTTGGGTACAGAACTGACACCAATCAAGGAGTGTAGAGG 551
 DB 1 ATGATTCACCATTCCTGCTTGGGTACAGAACTGACACCAATCAAGGAGTGTAGAGG 60
 QY 552 ACTGCTTCACAGCAGACCCCTGAGCTTACAGGGGCTGACGATGATGCTTCCCTTTC 611
 DB 61 ACTGCTTCACAGCAGACCCCTGAGCTTACAGGGGCTGACGATGATGCTTCCCTTTC 120
 QY 612 GCGCTGACAGGAAAGCGGCTTGTCTGCTCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 671
 DB 121 GCGCTGACAGGAAAGCGGCTTGTCTGCTCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 180
 QY 672 GTCTCCATCTACATCTCAACTGCTGCGGCGGACCTTCTCTTCTTACGGGCACTT 731
 DB 181 GTCTCCATCTACATCTCAACTGCTGCGGCGGACCTTCTCTTCTTACGGGCACTT 240
 QY 732 ATATGTTGGCGTTAGCGCTCATCAATATCCGCATCCCATCTCCAAATTCCTAGTCT 791
 DB 241 ATATGTTGGCGTTAGCGCTCATCAATATCCGCATCCCATCTCCAAATTCCTAGTCT 300
 QY 792 GTATGACCTTCCCTACTTTATAGGCTTAAGCATGCTAGAGGCTCATGACACCGAGCGC 851
 DB 301 GTATGACCTTCCCTACTTTATAGGCTTAAGCATGCTAGAGGCTCATGACACCGAGCGC 360
 QY 852 TGCCTGTCATCTGTGGCCCATCTGTGACCACTGCGCGGCCCATTAAGTCTGTATG 911

DB 361 TGCCTGTCATCTGTGGCCCATCTGTGACCACTGCGCGGCCCATTAAGTCTGTATG 420
 QY 912 GTATGACCTTCCCTACTTTATAGGCTTAAGCATGCTAGAGGCTCATGACACCGAGCGC 971
 DB 421 GTATGACCTTCCCTACTTTATAGGCTTAAGCATGCTAGAGGCTCATGACACCGAGCGC 480
 QY 972 TGTGACTTCTGTTTATAGTGTGCTGATCTGTTTGTGTGAAGTCAATTTCAATTACA 1031
 DB 481 TGTGACTTCTGTTTATAGTGTGCTGATCTGTTTGTGTGAAGTCAATTTCAATTACA 540
 QY 1032 ATGCGTGGCTGGTTTTTTTATAGTGTGCTGATCTGTTTGTGTGAAGTCAATTTCA 1091
 DB 541 ATGCGTGGCTGGTTTTTTTATAGTGTGCTGATCTGTTTGTGTGAAGTCAATTTCA 600
 QY 1092 AGGATTTCTGTGATGCCGGAAGATGCGGCTGACCAAGCTGTAGTGAACCATCTCTC 1151
 DB 601 AGGATTTCTGTGATGCCGGAAGATGCGGCTGACCAAGCTGTAGTGAACCATCTCTC 660
 QY 1152 ACAGTGTGCTTCTCTCTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 1211
 DB 661 ACAGTGTGCTTCTCTCTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 720
 QY 1212 AGGATCCACCTGGATGGAAGTCTTATTTGTATGTCATGATCTAGTTTCATTTCCCTG 1271
 DB 721 AGGATCCACCTGGATGGAAGTCTTATTTGTATGTCATGATCTAGTTTCATTTCCCTG 780
 QY 1272 TCCGCTCTTAAACAGAGTGCACACCCCATCTTACTTCTGAGGCTCTCTTTAGGCGAG 1331
 DB 781 TCCGCTCTTAAACAGAGTGCACACCCCATCTTACTTCTGAGGCTCTCTTTAGGCGAG 840
 QY 1332 CCGTAAATATAGCAGACACCTGAACTGTTCTCCAGAGGCTCTGACAGACACCGCTGAG 1391
 DB 841 CCGTAAATATAGCAGACACCTGAACTGTTCTCCAGAGGCTCTGACAGACACCGCTGAG 900
 QY 1392 GTGATGAAAGTGGAGGCTGCTCTGAGGAAACCCCTGAGCTGTGCGGAAGAGATTG 1451
 DB 901 GTGATGAAAGTGGAGGCTGCTCTGAGGAAACCCCTGAGCTGTGCGGAAGAGATTG 960
 QY 1452 GAGCAG 1457
 DB 961 GAGCAG 966

RESULT 12
 ID AAA29812
 AA29812 standard; cDNA; 969 BP.
 XX
 AC AAA29812;
 XX
 DT 18-AUG-2000 (first entry)
 XX
 DB Human G protein-coupled receptor hH7T213V encoding cDNA SEQ ID NO:4.
 XX
 KW Human; G protein-coupled receptor; hippocampus; diagnosis; screening;
 KW genetic disease; cellular function regulation; ss.
 XX
 OS Homo sapiens.
 XX
 FT Key Location/Qualifiers
 FT CDS 1..969
 FT /tag= a
 FT /product= "G protein-coupled receptor"
 XX
 PN MO20020455-A1.
 XX
 PD 13-APR-2000.
 XX
 PF 30-SEP-1999; 99WO-JP005366.
 XX
 PR 01-OCT-1998; 98UP-00279535.
 XX
 PA (TAKE) TAKEDA CHEM IND LTD.
 XX

CC useful for reducing intraocular pressure for the treatment of conditions
CC associated with elevated intraocular pressure such as glaucoma and
CC related conditions. The R-14 receptor, and host cells expressing an R-14
CC polynucleotide, may be used in screening for R-14 receptor antagonists.
CC The present sequence is related to the invention.

XX Sequence 969 BP; 171 A; 280 C; 250 G; 268 T; 0 U; 0 Other;

Query Match 35.0%; Score 916; DB 12; Length 969;
Best Local Similarity 99.9%; Pred. No. 0;
Matches 966; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

```
QY 492 ATGATTCACCATCCAGTCTTGAGTACAGAACTGACCAATCAACGAGAGAG 551
DB 1 ATGATTCACCATCCAGTCTTGAGTACAGAACTGACCAATCAACGAGAGAG 60
QY 552 ACTGCTTCAACAGACACCTGAGCTTCAAGGGGCTGACGTCATGCTTCCCTTGC 611
DB 61 ACTGCTTCAACAGACACCTGAGCTTCAAGGGGCTGACGTCATGCTTCCCTTGC 120
QY 612 GCGCTGACAGAGAAAGCGGCTTGTGCTGCTGCTGCTGCTGCTGCTGCTGCT 671
DB 121 GCGCTGACAGAGAGAGCGGCTTGTGCTGCTGCTGCTGCTGCTGCTGCTGCT 180
QY 672 GTCTTCATCAATCAATCTCAACTGCTGCTGCTGCTGCTGCTGCTGCTGCT 731
DB 181 GTCTTCATCAATCAATCTCAACTGCTGCTGCTGCTGCTGCTGCTGCTGCT 240
QY 732 ATATGTTGCGGCTTACGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 791
DB 241 ATATGTTGCGGCTTACGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 300
QY 792 GTGATGACCTTCCCTTCTTATAGAGCCATGAGCATGCTGAGGCGCATGACG 851
DB 301 GTGATGACCTTCCCTTCTTATAGAGCCATGAGCATGCTGAGGCGCATGACG 360
QY 852 TGCCCTGTCATCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 911
DB 361 TGCCCTGTCATCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 420
QY 912 GTCAATGTCATCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 971
DB 421 GTCAATGTCATCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 480
QY 972 TGTGATCTCTGTTAGTGTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 1031
DB 481 TGTGATCTCTGTTAGTGTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 540
QY 1032 ATGCGGTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 1091
DB 541 ATGCGGTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 600
QY 1092 AGGATTCCTGAGATCCCGAGAGATGCGGCTGACAGGCTGAGTGCATCCGCT 1151
DB 601 AGGATTCCTGAGATCCCGAGAGATGCGGCTGACAGGCTGAGTGCATCCGCT 660
QY 1152 ACAGTGTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 1211
DB 661 ACAGTGTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 720
QY 1212 AGATTCACCTGAGATGAGAACTTATTTTGTGATGCTGATGCTGATTTT 1271
DB 721 AGATTCACCTGAGATGAGAACTTATTTTGTGATGCTGATGCTGATTTT 780
QY 1272 TCCGCTTCAACAGAGTGCACACCCCATCTTCTTGTGAGGCTCCTTTAGGCG 1331
DB 781 TCCGCTTCAACAGAGTGCACACCCCATCTTCTTGTGAGGCTCCTTTAGGCG 840
QY 1332 CGTCAAAATAGGACAGAACTGAGTGTGCTTCAAGAGGCTCTGACAGACAG 1391
DB 841 CGTCAAAATAGGACAGAACTGAGTGTGCTTCCAGAGGCTCTGACAGACAG 900
QY 1392 GTGATGAAAGTGAAGGCTGCTGCTTCTCAAGAACTGAGAGCTGTGGGAA 1451
```

```
DB 901 GTGATGAAAGTGAAGGCTGCTTCTCAAGAACTGAGAGCTGTGGGAA 960
QY 1452 GAGCAGT 1458
DB 961 GAGCAGT 967
```

RESULT 14
AD044605
ID AD044605 standard; DNA; 966 BP.

AD044605;

29-JUL-2004 (first entry)

Human H177213 protein encoding DNA.

H177213, transgenic; G protein-coupled receptor; GPCR; ophthalmological;

cytostatic; nephrotoxic; antiinflammatory; dermatological; analgesic;

vulnerable; neuroprotective; human; gene; ds.

Homo sapiens.

Key Location/Qualifiers

FT CDS 1..966

FT FT /*tag= a

FT FT /product= "H177213"

MO200403972-A1.

13-MAY-2004.

28-OCT-2003; 2003WO-JP013781.

29-OCT-2002; 2002JP-00314141.

(TAKEDA) TAKEDA CHEM IND LTD.

Kaisho Y, Watanabe T, Yasuhara Y, Mori I, Taketomi S;

WPI: 2004-376191/35.

P-PSDB; AD044604.

H177213 protein, encoded DNA and transgenic animals for clarifying

pathological mechanism, developing therapeutic methods and screening

preventives or remedies for related diseases e.g. cataract, cancer, and

dermatitis.

disclosure; SEQ ID NO 4; 161bp; Japanese.

The invention relates to a non-human mammal that carries a DNA integrated

with a foreign H177213 or its mutant gene, or a part of it. The non-human

animal is particularly a rat. Such gene shows phenotypes of e.g. cataract

onset, transient skin rash and proliferation-promoting activity. The

foreign H177213 gene is a gene that encodes a G protein-coupled receptor

(GPCR) protein H177213. The protein, its encoded DNA and constructed

transgenic animals are useful for clarifying pathological mechanism,

developing therapeutic methods and screening preventives or remedies for

related diseases e.g. cataract, cancer, and dermatitis. The present

sequence represents a DNA encoding a human H177213 protein.

Sequence 966 BP; 171 A; 280 C; 248 G; 267 T; 0 U; 0 Other;

Query Match 35.0%; Score 915; DB 12; Length 966;
Best Local Similarity 99.9%; Pred. No. 0;
Matches 965; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

```
QY 492 ATGATTCACCATCCAGTCTTGAGTACAGAACTGACCAATCAACGAGAGAG 551
DB 1 ATGATTCACCATCCAGTCTTGAGTACAGAACTGACCAATCAACGAGAGAG 60
QY 552 ACTGCTTCAACAGACACCTGAGCTTCAAGGGGCTGACGTCATGCTTCCCTTGC 611
```


QY 852 TGCTGTCATCCTGTGGCCCATCTGTACACATGCCCCGCCCAAGATACCTGTGATCG 911
 Db |||||
 Db 361 TGCTGTCCATCTGTGGCCCATCTGTACACATGCCCCGCCCAAGATACCTGTGATCG 420
 QY 912 GTCAATGTGTCTGTGGCCCATCTGTACACATGCCCCGCCCAAGATACCTGTGATCG 971
 Db |||||
 Db 421 GTCAATGTGTCTGTGGCCCATCTGTACACATGCCCCGCCCAAGATACCTGTGATCG 480
 QY 972 TGTGACTTCTGTGTAGTGTGCTGATCTGTGTGTGTGAACGTCAAGATTTCAATTACA 1031
 Db |||||
 Db 481 TGTGACTTCTGTGTAGTGTGCTGATCTGTGTGTGTGAACGTCAAGATTTCAATTACA 540
 QY 1032 ATCCGCTGCTGTGTTTTTTATGTGTGTCTGTGTGTGTCAAGCTGTGCTGTGCTGTC 1091
 Db |||||
 Db 541 ATCCGCTGCTGTGTTTTTTATGTGTGTCTGTGTGTGTCAAGCTGTGCTGTGCTGTC 600
 QY 1092 AGGATTCCTGTGTATCCCGAAGATGCCGTGACCAAGCTGTATGATCCATCTCTCC 1151
 Db |||||
 Db 601 AGGATTCCTGTGTATCCCGAAGATGCCGTGACCAAGCTGTATGATCCATCTCTCC 660
 QY 1152 ACAGTGTGTCTTCTCTCTGTGTGCTGTGCTTGTGCTATCACTGTGCTGTGCTGTGCT 1211
 Db |||||
 Db 661 ACAGTGTGTCTTCTCTCTGTGTGCTGTGCTTGTGCTATCACTGTGCTGTGCTGTGCT 720
 QY 1212 AGGATTCACCTGTGTATGAAAGTCTTATTTGTCAATGTGATCTATGTTTCCATTTCTG 1271
 Db |||||
 Db 721 AGGATTCACCTGTGTATGAAAGTCTTATTTGTCAATGTGATCTATGTTTCCATTTCTG 780
 QY 1272 TCCGCTCTTAACAGAGTGCACACCCCATCAATTACTTGTGTGCTCTCTTTAGGCTG 1331
 Db |||||
 Db 781 TCCGCTCTTAACAGAGTGCACACCCCATCAATTACTTGTGTGCTCTCTTTAGGCTG 840
 QY 1332 CGTCAAAATAGGCAAGACCTGAGCTGTCTTCAAGAGGCTCTGACAGACAGCTGTGAG 1391
 Db |||||
 Db 841 CGTCAAAATAGGCAAGACCTGAGCTGTCTTCAAGAGGCTCTGACAGACAGCTGTGAG 900
 QY 1392 GTGATGAAGGTGAGGG 1409
 Db |||||
 Db 901 GTGATGAAGGTGAGGG 918

Search completed: October 29, 2004, 19:33:33
 Job time : 1232 secs

GenCore version 5.1.6
Copyright (c) 1993 - 2004 Compugen Ltd.

OM nucleic - nucleic search, using sw model

Run on: October 29, 2004, 18:52:43 ; Search time 215 seconds
(without alignments)

8655.094 Million cell updates/sec

Title: US-09-867-570-1

Perfect score: 2618

Sequence: 1 aacaatgcgcgcgaatcgcg.....gaaaaaaaaaaaaaaaaa 2618

Scoring table: OLIGO_NUC

Gapop 60.0 , Gapext 60.0

Searched: 824507 seqs, 355394441 residues

Word size : 20

Total number of hits satisfying chosen parameters: 2339

Minimum DB seq length: 0

Maximum DB seq length: 200000000

Post-processing: Listing first 45 summaries

Database : Issued Patents NA:*

- 1: /cgn2_6/ptodata/1/ina/5A.COMB.seq:*
- 2: /cgn2_6/ptodata/1/ina/5B.COMB.seq:*
- 3: /cgn2_6/ptodata/1/ina/6A.COMB.seq:*
- 4: /cgn2_6/ptodata/1/ina/6B.COMB.seq:*
- 5: /cgn2_6/ptodata/1/ina/PTCUS.COMB.seq:*
- 6: /cgn2_6/ptodata/1/ina/backfile1.seq:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	554	21.2	969	4 US-09-254-227A-6	Sequence 6, Appl1
2	546	20.9	969	4 US-09-254-227A-4	Sequence 4, Appl1
3	291	11.1	291	4 US-09-495-050A-164	Sequence 164, App
4	275	10.5	275	4 US-09-016-434-330	Sequence 330, App
5	179	6.8	969	4 US-09-254-227A-8	Sequence 8, Appl1
6	120	4.6	969	4 US-09-254-227A-10	Sequence 10, Appl1
7	80	3.1	969	4 US-09-254-227A-12	Sequence 12, Appl1
8	30	1.2	118067	4 US-09-497-855A-32	Sequence 32, Appl1
9	30	1.1	92139	4 US-09-918-686-1	Sequence 1, Appl1
10	26	1.0	735	3 US-08-871-732A-2	Sequence 2, Appl1
11	26	1.0	735	3 US-07-346-510B-2	Sequence 1, Appl1
12	26	1.0	682	1 US-07-972-481-1	Sequence 1, Appl1
13	24	0.9	1509	4 US-09-149-476-179	Sequence 179, App
14	24	0.9	87	4 US-09-621-976-14560	Sequence 14560, A
15	23	0.9	166	4 US-09-621-976-18390	Sequence 18390, A
16	23	0.9	166	4 US-09-702-705-112	Sequence 112, App
17	23	0.9	218	4 US-09-736-457-112	Sequence 112, App
18	23	0.9	218	4 US-09-614-124B-112	Sequence 112, App
19	23	0.9	218	4 US-09-671-325-112	Sequence 112, App
20	23	0.9	218	4 US-09-589-184-112	Sequence 112, App
21	23	0.9	218	4 US-09-658-824-112	Sequence 112, App
22	23	0.9	275	3 US-09-328-111-385	Sequence 385, App
23	23	0.9	357	4 US-09-513-999C-24748	Sequence 24748, A
24	23	0.9	479	4 US-09-621-976-893	Sequence 893, App
25	23	0.9	556	4 US-09-370-838-242	Sequence 242, App
26	23	0.9	556	4 US-09-854-133-242	Sequence 242, App
27	23	0.9	556	4 US-09-854-133-242	Sequence 242, App

28	23	0.9	651	4 US-09-641-638-72	Sequence 72, Appl1
29	23	0.9	651	4 US-10-170-097-72	Sequence 72, Appl1
30	23	0.9	761	4 US-09-641-638-71	Sequence 71, Appl1
31	23	0.9	761	4 US-10-170-097-71	Sequence 71, Appl1
32	23	0.9	947	4 US-09-673-395A-8	Sequence 8, Appl1
33	23	0.9	1092	4 US-09-461-325-86	Sequence 86, Appl1
34	23	0.9	1092	4 US-10-012-542-86	Sequence 86, Appl1
35	23	0.9	1092	4 US-10-115-123-86	Sequence 86, Appl1
36	23	0.9	1100	1 US-08-820-051-3	Sequence 3, Appl1
37	23	0.9	1126	4 US-09-461-325-89	Sequence 89, Appl1
38	23	0.9	1126	4 US-10-012-542-89	Sequence 89, Appl1
39	23	0.9	1126	4 US-10-115-123-89	Sequence 89, Appl1
40	23	0.9	1350	4 US-09-149-476-248	Sequence 248, App
41	23	0.9	1353	3 US-08-913-014A-6	Sequence 6, Appl1
42	23	0.9	1353	4 US-09-653-285-6	Sequence 6, Appl1
43	23	0.9	1364	4 US-09-828-995B-31	Sequence 31, Appl1
44	23	0.9	1364	4 US-09-828-995B-33	Sequence 33, Appl1
45	23	0.9	1425	1 US-08-464-148-1	Sequence 1, Appl1

ALIGNMENTS

RESULT 1					
US-09-254-227A-6					
; Sequence 6, Application US/09254227A					
; Patent No. 6696257					
; GENERAL INFORMATION:					
; APPLICANT: Ahmad, Sultan					
; APPLICANT: Barville, Denis					
; APPLICANT: Fortin, Yves					
; APPLICANT: Lembo, Paola					
; APPLICANT: O'Donnell, Dajan					
; APPLICANT: Shi-Hsiang, Shen					
; TITLE OF INVENTION: G Protein-Coupled Receptors from the Rat and Human					
; FILE REFERENCE: 81823/268117					
; CURRENT APPLICATION NUMBER: US/09/254, 227A					
; CURRENT FILING DATE: 1999-03-03					
; NUMBER OF SEQ ID NOS: 22					
; SOFTWARE: PatentIn version 3.0					
; SEQ ID NO 6					
; LENGTH: 969					
; TYPE: DNA					
; ORGANISM: Homo sapiens					
US-09-254-227A-6					
Query Match					
Best Local Similarity 99.4%; Pred. No. 3.9e-232;					
Matches 804; Conservative 0; Mismatches 5; Indels 0; Gaps 0;					
QY	505	TCCGAGTCTTGGGTACAGAACTGACACCAATCAACGACGTGAGAGACTCTTGCTTCA	564		
DB	14	TCCGAGTCTTGGGTACAGAACTGACACCAATCAACGACGTGAGAGACTCTTGCTTCA	73		
QY	565	AGCAGACCTTGAAGCTTCAACGAGGAGTGCATCGTTCCCTTGTGCGCTGACAGAA	624		
DB	74	AGCAGACCTTGAAGCTTCAACGAGGAGTGCATCGTTCCCTTGTGCGCTGACAGAA	133		
QY	625	ACGGGTTGTGCTGTGCTGCTGAGCTGCGCATGCGAGAACGCTGTCTCATTTACA	684		
DB	134	ACGGGTTGTGCTGTGCTGCTGAGCTGCGCATGCGAGAACGCTGTCTCATTTACA	193		
QY	685	TCTTCAACCTGCTGCGGCGGACCTTCTCTTCTTAAGGGGACATTAATATGTCGCGT	744		
DB	194	TCTTCAACCTGCTGCGGCGGACCTTCTCTTCTTAAGGGGACATTAATATGTCGCGT	253		
QY	745	TACGCTCATATATCGGATCCATCCATCTCAAAATCTCAGTCTGTGATGACCTTTC	804		
DB	254	TACGCTCATATATCGGATCCATCCATCTCAAAATCTCAGTCTGTGATGACCTTTC	313		
QY	805	CTTACTTATAGGCTTAAGCATGCTGAGCGGCATCAGACCGAGCGCTGCTGTCCATCC	864		
DB	314	CTTACTTATAGGCTTAAGCATGCTGAGCGGCATCAGACCGAGCGCTGCTGTCCATCC	373		

QY	865	GTGGCCCATCTGGATACCACTGCGCGGCCGACAGATACCTGATACGCGATAGTGTGCC	924
Db	374	TGTGCCCCATCTGGATACCACTGCGCGGCCGACAGATACCTGATACGCGATAGTGTGCC	433
QY	925	TGCTCTGGGCGCTGTTCCTCGTCCGCGAGATATCTCGAGTGGATGTTCTGTGACTTCTGT	984
Db	434	TGCTCTGGGCGCGCTGTTCCTCGTCCGCGAGATATCTCGAGTGGATGTTCTGTGACTTCTGT	493
QY	985	TTAGTGGTGTGATATCTGTGTTGGTGGAAACGTACGATTTCAATTCATACGAGTGGCTGG	1044
Db	494	TTAGTGGTGTGATATCTGTGTTGGTGGAAACGTACGATTTCAATTCATACGAGTGGCTGG	553
QY	1045	TTTTTTTATGTTGGTGTCTCTGTGGGTCCAGCGCTGTGCTCTGTGGTCAAGATTTCTGTGG	1104
Db	554	TTTTTTTATGTTGGTGTCTCTGTGGGTCCAGCGCTGTGCTCTGTGGTCAAGATTTCTGTGG	613
QY	1105	GATCCCGGAAGATGCGCTGACACAGAGCTGTACATCAATCTCTCCACACAGTGCAGTCT	1164
Db	614	GATCCCGGAAGATGCGCTGACACAGAGCTGTACATCAATCTCTCCACACAGTGCAGTCT	673
QY	1165	TCTCTCTCTGTGGCGCTGCTTGGCATTCAGTGGGCGCTGTGTTTCCAGATTCACACTGG	1224
Db	674	TCTCTCTCTGTGGCGCTGCTTGGCATTCAGTGGGCGCTGTGTTTCCAGATTCACACTGG	733
QY	1225	ATTGGAAGCTTTATTTTGTCAATGTGCATCTAGATTTCATATTTTCTGTCCGCTCTTAAACA	1284
Db	734	ATTGGAAGCTTTATTTTGTCAATGTGCATCTAGATTTCATATTTTCTGTCCGCTCTTAAACA	793
QY	1285	GCAGTGCCAACCCCATTCATTTACTCTTC	1313
Db	794	GCAGTGCCAACCCCATTCATTTACTCTTC	822

```

RESULT 2
US-09-254-227A-4
; Sequence 4, Application US/09254227A
; Patent No. 6696257
; GENERAL INFORMATION:
; APPLICANT: Ahmad, Sultan
; APPLICANT: Banville, Denis
; APPLICANT: Fortin, Yves
; APPLICANT: Lembo, Paola
; APPLICANT: O'Donnell, Dajan
; APPLICANT: Shi-Hsiang, Shen
; TITLE OF INVENTION: .G Protein-Coupled Receptors from the Rat and Human
; FILE REFERENCE: 81823/268117
; CURRENT APPLICATION NUMBER: US/09/254,227A
; CURRENT FILING DATE: 1999-03-03
; NUMBER OF SEQ ID NOS: 22
; SOFTWARE: PatentIn version 3.0
; SEQ ID NO 4
; LENGTH: 969
; TYPE: DNA
; ORGANISM: Homo sapiens
; US-09-254-227A-4

```

Query Match	20.9%	Score 546;	DB 4;	Length 969;
Best Local Similarity	99.8%;	Pred. No. 1.2e-228;		
Matches 596;	Conservative	0;	Mismatches 1;	Indels 0;
			Gaps	0

Qy	749	CCTCATCATATATCCGCATCCCACTCCGAAATTCCTCACTCCGATGATGACCTTCCCTCA	808
Db	258	CCTCATCAATATCCGCATCCCACTCCGAAATTCCTCACTCCGATGATGACCTTCCCTCA	317
Qy	809	CTTTATAGGCTTATAGCATGCTAGCGCATTCAGCACTCGAGCGCTGCCTTCATCCTGNG	868
Db	318	CTTTATAGGCTTATAGCATGCTAGCGCATTCAGCACTCGAGCGCTGCCTTCATCCTGNG	377
Qy	869	GGCCATCTGTGTACCATGCGCGCGCCCGAGATACCTGTATCGATCATGTGTGTCCTGTT	928
Db	378	GGCCATCTGTGTACCATGCGCGCGCCCGAGATACCTGTATCGATCATGTGTGTCCTGTT	437

Qy	929	CTGGGACCCTGTCCTCCGTGCGACGAAATATCCTGGAATGATGTTCTGTACATCTCCGTTAG	988
Dd	438	CTGGGACCCTGTCCTCCGTGCGGGAATATCCTGGATGGATGTTCTGTACATCTCCGTTAG	497
Qy	989	TGCTGCTGATTTCTGTGTTGGTGTGGAACGTCAGATTTCAATCAATCCGCTGCTGTTTT	1048
Dd	498	TGGTGTCTAATTTCTGTTGGTGTGGAACGTCAGATTTCAATCAATCCGCTGCTGTTTT	557
Qy	1049	TTTATATGCTGATTTCTGTGTGGATCAGACCTGATCTGTGGTCAGAAATTCCTGTGATC	1108
Dd	558	TTTATATGCTGATTTCTGTGTGGATCAGACCTGATCTGTGGTCAGAAATTCCTGTGATC	617
Qy	1109	CCGGAAGAATGCGCGTGACACAGGCTGTACGTAACATCCTCTCAACAGTGGTCTTCTCT	1168
Dd	618	CCGGAAGAATGCGCGTGACACAGGCTGTACGTAACATCCTCTCAACAGTGGTCTTCTCT	677
Qy	1169	CCTCTGTGGCCCTGCCCTTTGGCATTCAGATGGGCCCTGTGTTTCACAGATTCACCTGATG	1228
Dd	678	CCTCTGTGGCCCTGCCCTTTGGCATTCAGATGGGCCCTGTGTTTCACAGATTCACCTGATG	737
Qy	1229	GAAAGTCTTATTTTGTGATGTGCATCTAAGTTTCCATTTTCTGTGCCGCTCTTAACACAG	1288
Dd	738	GAAAGTCTTATTTTGTGATGTGCATCTAAGTTTCCATTTTCTGTGCCGCTCTTAACACAG	797
Qy	1289	TGCCAACCCTCAATTTACTTCTTCGTGGGCTCCCTTAAGCAGAGGTCAAAATAGCA	1345
Dd	798	TGCCAACCCTCAATTTACTTCTTCGTGGGCTCCCTTAAGCAGAGGTCAAAATAGCA	854

```

RESULT 3
US-09-495-050A-164
; Sequence 164, Application US/09495050A
; Patent No. 6492505
; GENERAL INFORMATION:
; APPLICANT: Roopa, Reddy
; APPLICANT: Guegler, Karl, J.
; APPLICANT: Au-Young, Janice
; TITLE OF INVENTION: COMPOSITION FOR DETECTION OF GENES ENCODING MEMBRANE-ASSOCIATED PROTEINS
; FILE REFERENCE: PA-0013 US
; CURRENT APPLICATION NUMBER: US/09/495,050A
; CURRENT FILING DATE: 2000-01-31
; PRIOR APPLICATION NUMBER: 60/118,318
; PRIOR FILING DATE: February 1, 1999
; NUMBER OF SEQ ID NOS: 305
; SOFTWARE: PERL Program
; SEQ ID NO 164
; LENGTH: 291
; TYPE: DNA
; ORGANISM: Homo sapiens
; FEATURES:
; NAME/KEY: misc feature
; OTHER INFORMATION: Incyte ID No. 6492505 1909132CT1
US-09-495-050A-164

```

Query Match	11.1%;	Score 291;	DB 4;	Length 291;
Best Local Similarity	100.0%;	Pred. No. 2.4e-117;		
Matches 291;	Conservative 0;	Mismatches 0;	Indels 0;	Gaps 0;

QY	762	GGCATCCCATCTCCAAAATCTAGTCCGTGTGATGACCTTCCCTACTTATAGGGCTA	821
Db	1	CGCATCCCATCTCCAAAATCTTAGTCCGTGTGATGACCTTCCCTACTTATAGGGCTA	60
QY	822	AGCATGCTGACGGCCATCAGCACCGAGCGCTGCTTCATCTGTGGCCATCTGTATAC	881
Db	61	AGCATGCTGACGGCCATCAGCACCGAGCGCTGCTTCATCTGTGGCCATCTGTATAC	120
QY	882	CATGCGCGCGCCCCAGATACCTGTCAATCGGTCAATGTGTGTCTGTGCTCTGGGCCCTGTCC	941
Db	121	CATGCGCGCGCCCCAGATACCTGTCAATCGGTCAATGTGTGTCTGTGCTCTGGGCCCTGTCC	180
QY	942	CTGCTGCGGAGATATCTGGAGTGTGATGTTCTGTATATCTTCGTTTATAGTGGTGTGATTC	1001
Db	181	CTGCTGCGGAGATATCTGGAGTGTGATGTTCTGTATATCTTCGTTTATAGTGGTGTGATTC	240

QY 1002 GTTGGTGAACGTCAGATTTCATTACATCGCGTGGCTGTTT 1052
DB 241 GTTGGTGAACGTCAGATTTCATTACATCGCGTGGCTGTTT 291

RESULT 4

US-09-016-434-330
Sequence 330, Application US/09016434
Patent No. 6500938
GENERAL INFORMATION:
APPLICANT: Janice Au-Young
APPLICANT: Jeffrey J. Sellhammer
TITLE OF INVENTION: COMPOSITION FOR THE DETECTION OF SIGNALING
TITLE OF INVENTION: PATHWAY GENE EXPRESSION
NUMBER OF SEQUENCES: 1490
CORRESPONDENCE ADDRESS:
ADDRESSEE: INCYTE PHARMACEUTICALS, INC.
STREET: 3174 PORTER DRIVE
CITY: PALO ALTO
STATE: CALIFORNIA
COUNTRY: USA
ZIP: 94304
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: Word Perfect 6.1 for Windows/MS-DOS 6.2
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/09/016.434
FILING DATE: HEREMITH
CLASSIFICATION:
PRIOR APPLICATION DATA:
APPLICATION NUMBER:
FILING DATE:
CLASSIFICATION:
ATTORNEY/AGENT INFORMATION:
NAME: Zeller, Karen J.
REGISTRATION NUMBER: 37,071
REFERENCE/DOCKET NUMBER: PA-0002 US
TELECOMMUNICATION INFORMATION:
TELEPHONE: (650) 855-0555
TELEFAX: (650) 845-4166
INFORMATION FOR SEQ ID NO: 330:
SEQUENCE CHARACTERISTICS:
LENGTH: 275 base pairs
TYPE: nucleic acid
STRANDEDNESS: single
TOPOLOGY: linear
IMMEDIATE SOURCE:
LIBRARY: CONNUT01
CLONE: 1909132
US-09-016-434-330

Query Match 10.5%; Score 275; DB 4; Length 275;
Best Local Similarity 100.0%; Pred. No. 2,3e-110;
Matches 275; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 762 CGCCATCCATCTTCAAAATCTCAGTCTGTGATGACCTTCCCTAATTATAGGCTTA 821
DB 1 CGCCATCCATCTTCAAAATCTCAGTCTGTGATGACCTTCCCTAATTATAGGCTTA 60
QY 822 AGCATGTGAGCGGCATCAGACCGAGCGCTGCTTCCATCTGTGGCCCATTTGTTAC 881
DB 61 AGCATGTGAGCGGCATCAGACCGAGCGCTGCTTCCATCTGTGGCCCATTTGTTAC 120
QY 882 CACTGCGCGCGCCCAAGTACCTGTCACTGGTCAATGTGCTGCTCTGGGCGCTGTC 941
DB 121 CACTGCGCGCGCCCAAGTACCTGTCACTGGTCAATGTGCTGCTCTGGGCGCTGTC 180
QY 942 CTGCTGGGAGATATCTTGAGATGATGTTCTGTGACTTCTGTTAGTGTGATTTCT 1001
DB 181 CTGCTGGGAGATATCTTGAGATGATGTTCTGTGACTTCTGTTAGTGTGATTTCT 240

QY 1002 GTTGGTGAACGTCAGATTTCATTACATCGC 1036
DB 241 GTTGGTGAACGTCAGATTTCATTACATCGC 275

RESULT 5

US-09-254-227A-8
Sequence 8, Application US/09254227A
Patent No. 6696257
GENERAL INFORMATION:
APPLICANT: Ahmad, Sultan
APPLICANT: Banville, Denis
APPLICANT: Fortin, Yves
APPLICANT: Lembo, Paola
APPLICANT: O'Donnell, Dajan
TITLE OF INVENTION: G Protein-Coupled Receptors from the Rat and Human
FILE REFERENCE: 81823/268117
CURRENT APPLICATION NUMBER: US/09/254.227A
CURRENT FILING DATE: 1999-03-03
NUMBER OF SEQ ID NOS: 22
SOFTWARE: PatentIn version 3.0
SEQ ID NO 8
LENGTH: 969
TYPE: DNA
ORGANISM: Homo sapiens
US-09-254-227A-8

Query Match 6.8%; Score 179; DB 4; Length 969;
Best Local Similarity 99.6%; Pred. No. 1.7e-68;
Matches 229; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 1231 AAGCTTATTTTGTGATGATGATCTTCCATTTTCTGCTCGCTTAAAGCAGTG 1290
DB 740 AAGCTTATTTTGTGATGATGATCTTCCATTTTCTGCTCGCTTAAAGCAGTG 799
QY 1291 CCAACCCATCATTTACTTCTTGCGGGCTCCTTTAGGAGCGCTCAAAATAGCAGAAC 1350
DB 800 CCAACCCATCATTTACTTCTTGCGGGCTCCTTTAGGAGCGCTCAAAATAGCAGAAC 859
QY 1351 TGAAGTGTGTTCTCCAGAGGGCTCTGAGAGACCGCTGAGGTGATGAAAGTGAAGGT 1410
DB 860 TGAAGTGTGTTCTCCAGAGGGCTCTGAGAGACCGCTGAGGTGATGAAAGTGAAGGT 919
QY 1411 GCGTTCTTCAGAAACCTTGAAGCTGTGCGGAAGCAGATTGAGCAGTGA 1460
DB 920 GCGTTCTTCAGAAACCTTGAAGCTGTGCGGAAGCAGATTGAGCAGTGA 969

RESULT 6

US-09-254-227A-10
Sequence 10, Application US/09254227A
Patent No. 6696257
GENERAL INFORMATION:
APPLICANT: Ahmad, Sultan
APPLICANT: Banville, Denis
APPLICANT: Fortin, Yves
APPLICANT: Lembo, Paola
APPLICANT: O'Donnell, Dajan
TITLE OF INVENTION: G Protein-Coupled Receptors from the Rat and Human
FILE REFERENCE: 81823/268117
CURRENT APPLICATION NUMBER: US/09/254.227A
CURRENT FILING DATE: 1999-03-03
NUMBER OF SEQ ID NOS: 22
SOFTWARE: PatentIn version 3.0
SEQ ID NO 10
LENGTH: 969
TYPE: DNA
ORGANISM: Homo sapiens
US-09-254-227A-10

Query Match 4.6%; Score 120; DB 4; Length 969;
Best Local Similarity 100.0%; Pred. No. 9.5e-43;
Matches 120; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1263 ATTTCTGTCGCTCTTAAACAGAGGCGCAACCCCATATTACTTCTTCTGGGCTCC 1322

DB 772 ATTTCCGTGCTCGCTCTTAAACAGAGGCGCAACCCCATATTACTTCTTCTGGGCTCC 831

QY 1323 TTATGCGAGCGCTCAAAATAGAGAGAACCTGAGCTGTTCTCAGAGGGCTCGAGAGAC 1382

DB 832 TTATGCGAGCGCTCAAAATAGAGAGAACCTGAGCTGTTCTCAGAGGGCTCGAGAGAC 891

RESULT 7
US-09-254-227A-12
; Sequence 12, Application US/09254227A

; Patent No. 6696257
; GENERAL INFORMATION:
; APPLICANT: Ahmad, Sultan

; APPLICANT: Barville, Denis

; APPLICANT: Fortin, Yves

; APPLICANT: Lembo, Paola

; APPLICANT: O'Donnell, Dajan

; APPLICANT: Shi-Hsiang, Shen

; TITLE OF INVENTION: G Protein-Coupled Receptors from the Rat and Human

; FILE REFERENCE: 81823/268117

; CURRENT APPLICATION NUMBER: US/09/254,227A

; CURRENT FILING DATE: 1999-03-03

; NUMBER OF SEQ ID NOS: 22

; SOFTWARE: Patentin version 3.0

; SEQ ID NO 12

; LENGTH: 969

; TYPE: DNA

; ORGANISM: Homo sapiens

US-09-254-227A-12

Query Match 3.7%; Score 97; DB 4; Length 969;
Best Local Similarity 100.0%; Pred. No. 1e-32;
Matches 97; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1287 AGTGGCAACCCCATATTACTTCTTCTGGGCTCTTTAGGAGGCTCAAAATAGGCGAG 1346

DB 796 AGTGGCAACCCCATATTACTTCTTCTGGGCTCTTTAGGAGGCTCAAAATAGGCGAG 855

QY 1347 AACCTGAAGCTGTTCTCAGAGGGCTCGAGAGACA 1383

DB 856 AACCTGAAGCTGTTCTCAGAGGGCTCGAGAGACA 892

RESULT 8
US-09-254-227A-14
; Sequence 14, Application US/09254227A

; Patent No. 6696257
; GENERAL INFORMATION:
; APPLICANT: Ahmad, Sultan

; APPLICANT: Barville, Denis

; APPLICANT: Fortin, Yves

; APPLICANT: Lembo, Paola

; APPLICANT: O'Donnell, Dajan

; APPLICANT: Shi-Hsiang, Shen

; TITLE OF INVENTION: G Protein-Coupled Receptors from the Rat and Human

; FILE REFERENCE: 81823/268117

; CURRENT APPLICATION NUMBER: US/09/254,227A

; CURRENT FILING DATE: 1999-03-03

; NUMBER OF SEQ ID NOS: 22

; SOFTWARE: Patentin version 3.0

; SEQ ID NO 14

; LENGTH: 969

; TYPE: DNA

; ORGANISM: Homo sapiens

US-09-254-227A-14

Query Match 3.1%; Score 80; DB 4; Length 969;

Best Local Similarity 100.0%; Pred. No. 2.7e-25;
Matches 80; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1287 AGTGGCAACCCCATATTACTTCTGAGGCTCTTTAGGAGGCTCAAAATAGGCGAG 1346

DB 796 AGTGGCAACCCCATATTACTTCTGAGGCTCTTTAGGAGGCTCAAAATAGGCGAG 855

QY 1347 AACCTGAAGCTGTTCTCCA 1366

DB 856 AACCTGAAGCTGTTCTCCA 875

RESULT 9
US-09-497-855A-32/c
; Sequence 32, Application US/09497855A

; Patent No. 6605432
; GENERAL INFORMATION:
; APPLICANT: Huang, Tim

; TITLE OF INVENTION: HIGH-THROUGHPUT METHODS FOR DETECTING DNA METHYLATION

; FILE REFERENCE: UMO1523

; CURRENT APPLICATION NUMBER: US/09/497,855A

; CURRENT FILING DATE: 2000-02-04

; PRIOR APPLICATION NUMBER: 60/120,592

; PRIOR FILING DATE: 1999-02-18

; PRIOR APPLICATION NUMBER: 60/118,760

; PRIOR FILING DATE: 1999-02-05

; NUMBER OF SEQ ID NOS: 54

; SOFTWARE: Patentin version 3.0

; SEQ ID NO 32

; LENGTH: 118067

; TYPE: DNA

; ORGANISM: Homo sapiens;

US-09-497-855A-32

Query Match 1.2%; Score 31; DB 4; Length 118067;
Best Local Similarity 100.0%; Pred. No. 0.00051;
Matches 31; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 127 GATTGAAGCTTCTGAGGCTCCCGCAGAG 157

DB 97535 GATTGAAGCTTCTGAGGCTCCCGCAGAG 97505

RESULT 10
US-09-918-686-1/c
; Sequence 1, Application US/09918686

; Patent No. 6475739
; GENERAL INFORMATION:
; APPLICANT: Brunkow, Mary

; APPLICANT: Paepfer, Bryan

; APPLICANT: Steahling-Hampton, Karen

; TITLE OF INVENTION: GENOMIC DELETIONS

; FILE REFERENCE: 240083.515

; CURRENT APPLICATION NUMBER: US/09/918,686

; CURRENT FILING DATE: 2001-07-30

; NUMBER OF SEQ ID NOS: 105

; SOFTWARE: FastSeq for Windows Version 4.0

; SEQ ID NO 1

; LENGTH: 92139

; TYPE: DNA

; ORGANISM: Homo sapiens

; FEATURE: NAME/KEY: misc_feature

; LOCATION: 7043, 8369, 8401

; OTHER INFORMATION: n = A,T,C or G

US-09-918-686-1

Query Match 1.1%; Score 30; DB 4; Length 92139;
Best Local Similarity 100.0%; Pred. No. 0.0014;
Matches 30; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 127 GATTGGAAGCTTCTGAGGCTCCCGCAGAA 156
|||||
Db 80505 GATTGGAAGCTTCTGAGGCTCCCGCAGAA 80476

RESULT 11

US-08-871-732A-2
; Sequence 2, Application US/08871732A
; Patent No. 6140074
; GENERAL INFORMATION:
; APPLICANT: O'BRIEN, TIMOTHY J.
; TITLE OF INVENTION: NOVEL SH3 PROTEIN, GENE, CHIMERIC
; TITLE OF INVENTION: CELLS, VECTORS AND EXPRESSION METHOD FOR PRODUCING THE NOVEL
; TITLE OF INVENTION: PROTEIN, ANTIBODIES AND USES
; NUMBER OF SEQUENCES: 16
; CORRESPONDENCE ADDRESSES:
; ADDRESSEE: MARTIN L. MCGREGOR
; STREET: 5380 WEST 34TH STREET, #345
; CITY: HOUSTON
; STATE: TEXAS
; COUNTRY: UNITED STATES OF AMERICA
; ZIP: 77092
; COMPUTER READABLE FORM:
; MEDIUM TYPE: DISKETTE 3.5 INCH 1.44 MB STORAGE
; COMPUTER: IBM COMPATIBLE
; OPERATING SYSTEM: MS-DOS
; SOFTWARE: WORDPERFECT 6.1
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/871,732A
; FILING DATE:
; CLASSIFICATION: 435
; PRIOR APPLICATION DATA:
; ATTORNEY/AGENT INFORMATION:
; NAME: MCGREGOR, MARTIN L.
; REGISTRATION NUMBER: 29,329
; REFERENCE/DOCKET NUMBER: 1-1
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 713-682-1213
; TELEFAX: 713-682-5807
; TELEX: NONE
; INFORMATION FOR SEQ ID NO: 2:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 735 BASE PAIRS
; TYPE: NUCLEIC ACID
; STRANDEDNESS: SINGLE
; TOPOLOGY: LINEAR
; MOLECULE TYPE: OTHER NUCLEIC ACID
; HYPOTHEICAL: NO
; ANTI-SENSE: NO
US-08-871-732A-2

Query Match 1.0%; Score 26; DB 3; Length 735;
Best Local Similarity 100.0%; Pred. No. 0.1;
Matches 26; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 2593 GTGCTGAAAAA 2618
|||||
Db 709 GTGCTGAAAAA 734

RESULT 12
US-09-346-510B-2
; Sequence 2, Application US/09346510B
; Patent No. 6281014
; GENERAL INFORMATION:
; APPLICANT: O'Brien, Timothy J.
; TITLE OF INVENTION: SH3-Containing Protein, DNA and Uses Thereof
; FILE REFERENCE: D6221CIP
; CURRENT APPLICATION NUMBER: US/09/346,510B
; CURRENT FILING DATE: 1999-07-01
; PRIOR APPLICATION NUMBER: 08/871,732

; PRIOR FILING DATE: 1997-06-09
; NUMBER OF SEQ ID NOS: 32
; SEQ ID NO 2
; LENGTH: 735
; TYPE: DNA
; ORGANISM: Homo sapiens
; FEATURE:
; OTHER INFORMATION: cDNA sequence of TADG5 gene isolated from a cDNA
US-09-346-510B-2

Query Match 1.0%; Score 26; DB 3; Length 735;
Best Local Similarity 100.0%; Pred. No. 0.1;
Matches 26; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 2593 GTGCTGAAAAA 2618
|||||
Db 709 GTGCTGAAAAA 734

RESULT 13

US-07-972-481-1
; Sequence 1, Application US/07972481
; Patent No. 5578453
; GENERAL INFORMATION:
; APPLICANT: MCDONALD, PETER J
; TITLE OF INVENTION: CLONING AND EXPRESSION OF TOXOPLASMA ANTIGENS AND
; TITLE OF INVENTION: USE OF RECOMBINANT ANTIGENS
; NUMBER OF SEQUENCES: 4
; CORRESPONDENCE ADDRESSES:
; ADDRESSEE: BIRCH, STEWART, KOLASCH & BIRCH
; STREET: 8110 GATEHOUSE ROAD
; CITY: FALLS CHURCH
; STATE: VIRGINIA
; COUNTRY: USA
; ZIP: 22040-0747
; COMPUTER READABLE FORM:
; MEDIUM TYPE: FLOPPY disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.25
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/07/972,481
; FILING DATE: 12-APR-1993
; CLASSIFICATION: 435
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: INTERNATIONAL PCT/AU91/00347
; FILING DATE: 09-AUG-1991
; ATTORNEY/AGENT INFORMATION:
; NAME: SVENSSON, LEONARD R
; REGISTRATION NUMBER: 30,350
; REFERENCE/DOCKET NUMBER: 446-102PCT
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (703) 205-8000
; TELEFAX: (703) 205 8050
; INFORMATION FOR SEQ ID NO: 1:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 682 base pairs
; TYPE: nucleic acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; MOLECULE TYPE: DNA (genomic)
; FEATURE:
; NAME/KEY: CDS
; LOCATION: 1..289
US-07-972-481-1

Query Match 0.9%; Score 24; DB 1; Length 682;
Best Local Similarity 100.0%; Pred. No. 0.75;
Matches 24; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 2595 GTGCTGAAAAA 2618

DB 650 GCTGAAAAAAAAAAAAAAAAAAAA 673

RESULT 14
US-09-149-476-179

Sequence 179, Application US/09149476

Patent No. 6420526

GENERAL INFORMATION:

APPLICANT: Rosen et al.

TITLE OF INVENTION: 186 Human Secreted proteins

FILE REFERENCE: P2002p1

CURRENT APPLICATION NUMBER: US/09/149,476

EARLIER FILING DATE: 1998-09-08

EARLIER APPLICATION NUMBER: PCT/US98/04493

EARLIER FILING DATE: 1998-03-06

EARLIER APPLICATION NUMBER: 60/040,162

EARLIER FILING DATE: 1997-03-07

EARLIER APPLICATION NUMBER: 60/040,333

EARLIER FILING DATE: 1997-03-07

EARLIER APPLICATION NUMBER: 60/038,621

EARLIER FILING DATE: 1997-03-07

EARLIER APPLICATION NUMBER: 60/040,626

EARLIER FILING DATE: 1997-03-07

EARLIER APPLICATION NUMBER: 60/040,334

EARLIER FILING DATE: 1997-03-07

EARLIER APPLICATION NUMBER: 60/040,336

EARLIER FILING DATE: 1997-03-07

EARLIER APPLICATION NUMBER: 60/040,163

EARLIER FILING DATE: 1997-03-07

EARLIER APPLICATION NUMBER: 60/047,600

EARLIER FILING DATE: 1997-05-23

EARLIER APPLICATION NUMBER: 60/047,615

EARLIER FILING DATE: 1997-05-23

EARLIER APPLICATION NUMBER: 60/047,597

EARLIER FILING DATE: 1997-05-23

EARLIER APPLICATION NUMBER: 60/047,502

EARLIER FILING DATE: 1997-05-23

EARLIER APPLICATION NUMBER: 60/047,633

EARLIER FILING DATE: 1997-05-23

EARLIER APPLICATION NUMBER: 60/047,583

EARLIER FILING DATE: 1997-05-23

EARLIER APPLICATION NUMBER: 60/047,583

EARLIER FILING DATE: 1997-05-23

EARLIER APPLICATION NUMBER: 60/047,617

EARLIER FILING DATE: 1997-05-23

EARLIER APPLICATION NUMBER: 60/047,618

EARLIER FILING DATE: 1997-05-23

EARLIER APPLICATION NUMBER: 60/047,503

EARLIER FILING DATE: 1997-05-23

EARLIER APPLICATION NUMBER: 60/047,592

EARLIER FILING DATE: 1997-05-23

EARLIER APPLICATION NUMBER: 60/047,581

EARLIER FILING DATE: 1997-05-23

EARLIER APPLICATION NUMBER: 60/047,584

EARLIER FILING DATE: 1997-05-23

EARLIER APPLICATION NUMBER: 60/047,500

EARLIER FILING DATE: 1997-05-23

EARLIER APPLICATION NUMBER: 60/047,587

EARLIER FILING DATE: 1997-05-23

EARLIER APPLICATION NUMBER: 60/047,492

EARLIER FILING DATE: 1997-05-23

EARLIER APPLICATION NUMBER: 60/047,598

EARLIER FILING DATE: 1997-05-23

EARLIER APPLICATION NUMBER: 60/047,613

EARLIER FILING DATE: 1997-05-23

EARLIER APPLICATION NUMBER: 60/047,582

EARLIER FILING DATE: 1997-05-23

EARLIER APPLICATION NUMBER: 60/047,596

EARLIER FILING DATE: 1997-05-23

EARLIER APPLICATION NUMBER: 60/047,612

EARLIER FILING DATE: 1997-05-23

EARLIER APPLICATION NUMBER: 60/047,632

EARLIER FILING DATE: 1997-05-23

EARLIER APPLICATION NUMBER: 60/047,601

EARLIER FILING DATE: 1997-05-23

EARLIER APPLICATION NUMBER: 60/043,580

EARLIER FILING DATE: 1997-04-11

EARLIER APPLICATION NUMBER: 60/043,568

EARLIER FILING DATE: 1997-04-11

EARLIER APPLICATION NUMBER: 60/043,314

EARLIER FILING DATE: 1997-04-11

EARLIER APPLICATION NUMBER: 60/043,569

EARLIER FILING DATE: 1997-04-11

EARLIER APPLICATION NUMBER: 60/043,311

EARLIER FILING DATE: 1997-04-11

EARLIER APPLICATION NUMBER: 60/043,671

EARLIER FILING DATE: 1997-04-11

EARLIER APPLICATION NUMBER: 60/043,674

EARLIER FILING DATE: 1997-04-11

EARLIER APPLICATION NUMBER: 60/043,669

EARLIER FILING DATE: 1997-04-11

EARLIER APPLICATION NUMBER: 60/043,312

EARLIER FILING DATE: 1997-04-11

EARLIER APPLICATION NUMBER: 60/043,313

EARLIER FILING DATE: 1997-04-11

EARLIER APPLICATION NUMBER: 60/043,672

EARLIER FILING DATE: 1997-04-11

EARLIER APPLICATION NUMBER: 60/043,315

EARLIER FILING DATE: 1997-04-11

EARLIER APPLICATION NUMBER: 60/048,974

EARLIER FILING DATE: 1997-06-06

EARLIER APPLICATION NUMBER: 60/056,886

EARLIER FILING DATE: 1997-08-22

EARLIER APPLICATION NUMBER: 60/056,877

EARLIER FILING DATE: 1997-08-22

EARLIER APPLICATION NUMBER: 60/056,889

EARLIER FILING DATE: 1997-08-22

EARLIER APPLICATION NUMBER: 60/056,893

EARLIER FILING DATE: 1997-08-22

EARLIER APPLICATION NUMBER: 60/056,630

EARLIER FILING DATE: 1997-08-22

EARLIER APPLICATION NUMBER: 60/056,878

EARLIER FILING DATE: 1997-08-22

EARLIER APPLICATION NUMBER: 60/056,662

EARLIER FILING DATE: 1997-08-22

EARLIER APPLICATION NUMBER: 60/056,872

EARLIER FILING DATE: 1997-08-22

EARLIER APPLICATION NUMBER: 60/056,882

EARLIER FILING DATE: 1997-08-22

EARLIER APPLICATION NUMBER: 60/056,894

EARLIER FILING DATE: 1997-08-22

EARLIER APPLICATION NUMBER: 60/056,911

EARLIER FILING DATE: 1997-08-22

EARLIER APPLICATION NUMBER: 60/056,636

EARLIER FILING DATE: 1997-08-22

EARLIER APPLICATION NUMBER: 60/056,874

EARLIER FILING DATE: 1997-08-22

EARLIER APPLICATION NUMBER: 60/056,910

EARLIER FILING DATE: 1997-08-22

EARLIER APPLICATION NUMBER: 60/056,864

EARLIER FILING DATE: 1997-08-22

EARLIER APPLICATION NUMBER: 60/056,631

EARLIER FILING DATE: 1997-08-22

EARLIER APPLICATION NUMBER: 60/056,845

EARLIER FILING DATE: 1997-08-22

EARLIER APPLICATION NUMBER: 60/056,892

EARLIER FILING DATE: 1997-08-22

EARLIER APPLICATION NUMBER: 60/056,892

EARLIER FILING DATE: 1997-08-22

```

; EARLIER APPLICATION NUMBER: 60/057,761
; EARLIER FILING DATE: 1997-08-22
; EARLIER APPLICATION NUMBER: 60/047,595
; EARLIER FILING DATE: 1997-05-23
; EARLIER APPLICATION NUMBER: 60/047,599
; EARLIER FILING DATE: 1997-05-23
; EARLIER APPLICATION NUMBER: 60/047,588
; EARLIER FILING DATE: 1997-05-23
; EARLIER APPLICATION NUMBER: 60/047,585
; EARLIER FILING DATE: 1997-05-23
; EARLIER APPLICATION NUMBER: 60/047,586
; EARLIER FILING DATE: 1997-05-23
; EARLIER APPLICATION NUMBER: 60/047,590
; EARLIER FILING DATE: 1997-05-23
; EARLIER APPLICATION NUMBER: 60/047,594
; EARLIER FILING DATE: 1997-05-23
; EARLIER APPLICATION NUMBER: 60/047,589
; EARLIER FILING DATE: 1997-05-23
; EARLIER APPLICATION NUMBER: 60/047,593
; EARLIER FILING DATE: 1997-05-23
; EARLIER APPLICATION NUMBER: 60/047,614
; EARLIER FILING DATE: 1997-05-23
; EARLIER APPLICATION NUMBER: 60/043,578
; EARLIER FILING DATE: 1997-04-11
; EARLIER APPLICATION NUMBER: 60/043,576
; EARLIER FILING DATE: 1997-04-11
; EARLIER APPLICATION NUMBER: 60/047,501
; EARLIER FILING DATE: 1997-05-23
; EARLIER APPLICATION NUMBER: 60/043,670
; EARLIER FILING DATE: 1997-04-11
; EARLIER APPLICATION NUMBER: 60/056,632
; EARLIER FILING DATE: 1997-08-22
; EARLIER APPLICATION NUMBER: 60/056,664
; EARLIER FILING DATE: 1997-08-22
; EARLIER APPLICATION NUMBER: 60/056,876
; EARLIER FILING DATE: 1997-08-22
; EARLIER APPLICATION NUMBER: 60/056,881
; EARLIER FILING DATE: 1997-08-22
; EARLIER APPLICATION NUMBER: 60/056,909
; EARLIER FILING DATE: 1997-08-22
; EARLIER APPLICATION NUMBER: 60/056,875
; EARLIER FILING DATE: 1997-08-22
; EARLIER APPLICATION NUMBER: 60/056,862
; EARLIER FILING DATE: 1997-08-22
; EARLIER APPLICATION NUMBER: 60/056,887
; EARLIER FILING DATE: 1997-08-22
; EARLIER APPLICATION NUMBER: 60/056,908
; EARLIER FILING DATE: 1997-08-22
; EARLIER APPLICATION NUMBER: 60/048,964
; EARLIER FILING DATE: 1997-06-06
; EARLIER APPLICATION NUMBER: 60/057,650
; EARLIER FILING DATE: 1997-09-05
; EARLIER APPLICATION NUMBER: 60/056,884
; EARLIER FILING DATE: 1997-08-22
; EARLIER APPLICATION NUMBER: 60/057,669
; EARLIER FILING DATE: 1997-09-05
; EARLIER APPLICATION NUMBER: 60/049,610
; EARLIER FILING DATE: 1997-06-13
; EARLIER APPLICATION NUMBER: 60/061,060
; EARLIER FILING DATE: 1997-10-02

```

```

Query Match 0.9%; Score 24; DB 4; Length 1509;
Best Local Similarity 100.0%; Pred. No. 0.72;
Matches 24; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

```

```

QY 2595 GCTGAAAAAAAAAAAAAAAAAAAA 2618
DB 1428 GCTGAAAAAAAAAAAAAAAAAAAA 1451

```

```

RESULT 15
US-09-621-976-14560
; Sequence 14560, Application US/09621976

```

```

; Patent No. 6639063
; GENERAL INFORMATION:
; APPLICANT: Dumas Milne Edwards, J.B.
; APPLICANT: Jobert, S.
; APPLICANT: Giordano, J.Y.
; TITLE OF INVENTION: ESTs and Encoded Human Proteins.
; FILE REFERENCE: GENSET.054PR2
; CURRENT APPLICATION NUMBER: US/09/621,976
; CURRENT FILING DATE: 2000-07-21
; NUMBER OF SEQ ID NOS: 19335
; SOFTWARE: Patent.pm
; SEQ ID NO 14560
; LENGTH: 87
; TYPE: DNA
; ORGANISM: Homo sapiens
US-09-621-976-14560

```

```

Query Match 0.9%; Score 23; DB 4; Length 87;
Best Local Similarity 100.0%; Pred. No. 2.3;
Matches 23; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

```

```

QY 2596 CTGGAIAAAAAAAAAAAAAAAAAA 2618
DB 32 CTGGAIAAAAAAAAAAAAAAAAAA 54

```

```

Search completed: October 30, 2004, 00:50:23
Job time : 221 secs

```

This Page Blank (uspto)


```
Db 121 CACAGAGATTGGAACTTCCGAGGCTCCAGAAAGAGAGCTGATGCTTCTTGA 180
Qy 181 CAGTCTGAGAGCTTATGACGATTAAACCAATTCCTTCATTAATTTCCAGTCTCAG 240
Db 181 CAGTCTGAGAGCTTATGACGATTAAACCAATTCCTTCATTAATTTCCAGTCTCAG 240
Qy 241 TATTTCTTTTGAAGATTTGAGAAATGAATCAAGACAGAGAGCCAGAGATGAA 300
Db 241 TATTTCTTTTGAAGATTTGAGAAATGAATCAAGACAGAGAGCCAGAGATGAA 300
Qy 301 ATCCCAAGGTCTTCTGCTCTTCCAGTCTCTGCTGCTGCTGCTGCTGCTGCTGCT 360
Db 301 ATCCCAAGGTCTTCTGCTCTTCCAGTCTCTGCTGCTGCTGCTGCTGCTGCTGCT 360
Qy 361 TCACACAGAAACCAAGAAATGAAATCCCACTGATGCTGATGATGATGATGATGATG 420
Db 361 TCACACAGAAACCAAGAAATGAAATCCCACTGATGCTGATGATGATGATGATGATG 420
Qy 421 TGGGATGTCAAACAGATTAAGAAATGAAAGCAAAATCTCATGAGTCAATCAAGCTG 480
Db 421 TGGGATGTCAAACAGATTAAGAAATGAAAGCAAAATCTCATGAGTCAATCAAGCTG 480
Qy 481 GCTTTCTGAGATGATTCACATCCAGTCTTGGGTGACAACTGAAACCAATCAAG 540
Db 481 GCTTTCTGAGATGATTCACATCCAGTCTTGGGTGACAACTGAAACCAATCAAG 540
Qy 541 GACGTGAGAGAGCTCTTGTCTCAAGCAGACCTGAGCTTCAAGGGGCTGAGTGCATG 600
Db 541 GACGTGAGAGAGCTCTTGTCTCAAGCAGACCTGAGCTTCAAGGGGCTGAGTGCATG 600
Qy 601 TTTCTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 660
Db 601 TTTCTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 660
Qy 661 GCAGAAAGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 720
Db 661 GCAGAAAGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 720
Qy 721 GCGGACCATTAATATGTTGCGCGTTACGCTCATCAATATCCGCCATCCCATCCAAA 780
Db 721 GCGGACCATTAATATGTTGCGCGTTACGCTCATCAATATCCGCCATCCCATCCAAA 780
Qy 781 TCTCTAGTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 840
Db 781 TCTCTAGTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 840
Qy 841 GCACCGAGGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 900
Db 841 GCACCGAGGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 900
Qy 901 ACCTGTCAATCGGTCAATGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 960
Db 901 ACCTGTCAATCGGTCAATGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 960
Qy 961 AGTGAATGTTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 1020
Db 961 AGTGAATGTTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 1020
Qy 1021 ATTTCAATTAACATGCGGTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 1080
Db 1021 ATTTCAATTAACATGCGGTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 1080
Qy 1081 TCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 1140
Db 1081 TCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 1140
Qy 1141 CCATCTCTCTCAAGATGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 1200
Db 1141 CCATCTCTCTCAAGATGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 1200
Qy 1201 CCGTCTTTTCCAGATCCAGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 1260
Db 1201 CCGTCTTTTCCAGATCCAGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 1260

Db 1201 CCGTCTTTTCCAGATCCAGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 1260
Qy 1261 CCATTTTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 1320
Db 1261 CCATTTTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 1320
Qy 1321 CTTTGAAGCAGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 1380
Db 1321 CTTTGAAGCAGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 1380
Qy 1381 ACAGCCTGAGTGAATGAGTGAAGGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 1440
Db 1381 ACAGCCTGAGTGAATGAGTGAAGGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 1440
Qy 1441 GAAGCAGATTTGAGCAGTGAAGAAACCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 1500
Db 1441 GAAGCAGATTTGAGCAGTGAAGAAACCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 1500
Qy 1501 TGTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 1560
Db 1501 TGTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 1560
Qy 1561 TCTGAGTGTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 1620
Db 1561 TCTGAGTGTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 1620
Qy 1621 CACCATGAGAAAGCATTAATGCTGACAGTACATGCTGCTGCTGCTGCTGCTGCTGCT 1680
Db 1621 CACCATGAGAAAGCATTAATGCTGACAGTACATGCTGCTGCTGCTGCTGCTGCTGCT 1680
Qy 1681 ACATTTTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 1740
Db 1681 ACATTTTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 1740
Qy 1741 ATTGTAATTAAGAGTGTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 1800
Db 1741 ATTGTAATTAAGAGTGTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 1800
Qy 1801 CCTGATGATTAACAAAGAAAGATTCCTTAATATCTGCTGCTGCTGCTGCTGCTGCTGCT 1860
Db 1801 CCTGATGATTAACAAAGAAAGATTCCTTAATATCTGCTGCTGCTGCTGCTGCTGCTGCT 1860
Qy 1861 TCAATTTTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 1920
Db 1861 TCAATTTTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 1920
Qy 1921 GAGTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 1980
Db 1921 GAGTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 1980
Qy 1981 GAGAGAGATTCCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 2040
Db 1981 GAGAGAGATTCCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 2040
Qy 2041 CTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 2100
Db 2041 CTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 2100
Qy 2101 GCGCTTCCCTTGAAGACTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 2160
Db 2101 GCGCTTCCCTTGAAGACTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 2160
Qy 2161 TGAGAGGCTTAATTAATGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 2220
Db 2161 TGAGAGGCTTAATTAATGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 2220
Qy 2221 ATTTTAAGTGAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAAT 2280
Db 2221 ATTTTAAGTGAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAAT 2280
Qy 2281 GTAGAGATTCCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 2340
Db 2281 GTAGAGATTCCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 2340
```


QY	2341	TAGTTTATGGCTCTTTACATGAGAAGCAAACTGAACTTTCATGACAA	2400
QY	2341	TAGTTTATGGCTCTTTACATGAGAAGCAAACTGAACTTTCATGACAA <td>2400</td>	2400
Db	2341	TAGTTTATGGCTCTTTACATGAGAAGCAAACTGAACTTTCATGACAA	2400
QY	2401	TCTCAATTATGTCATCTGAGTAATPACTTACAGTTGGTACAGAAATCTGATACATGCTGT	2460
Db	2401	TCTCAATTATGTCATCTGAGTAATPACTTACAGTTGGTACAGAAATCTGATACATGCTGT	2460
QY	2461	GACATACATGAACCTGGAAATATTGTGCTTAAGAAAAATPACCGCAAGCCCAAACTATTT	2520
Db	2461	GACATACATGAACCTGGAAATATTGTGCTTAAGAAAAATPACCGCAAGCCCAAACTATTT	2520
QY	2521	GTAAGTTCAAAATCTATGAGGTATCCAAATTAGGAATCTTTGAACACAGAAAAATAAAT	2580
Db	2521	GTAAGTTCAAAATCTATGAGGTATCCAAATTAGGAATCTTTGAACACAGAAAAATAAAT	2580
QY	2581	AGGAGATCCTGTGCTGCAAAAAAAAAAAAAAAAAAAAAA	2618
Db	2581	AGGAGATCCTGTGCTGCAAAAAAAAAAAAAAAAAAAAAA	2618

```

RESULT 2
US-09-867-570-3
; Sequence 3, Application US/09867570
; Publication No. US20040076951A1
; GENERAL INFORMATION:
; APPLICANT: WEI, Ming-Hui et al.
; TITLE OF INVENTION: ISOLATED HUMAN G-PROTEIN COUPLED
; TITLE OF INVENTION: RECEPTORS, NUCLEIC ACID MOLECULES ENCODING HUMAN GPCR
; TITLE OF INVENTION: PROTEINS, AND USBS THEREOF
; FILE REFERENCE: CL000900-CIP
; CURRENT APPLICATION NUMBER: US/09/867,570
; CURRENT FILING DATE: 2001-05-31
; PRIOR APPLICATION NUMBER: 09/695,045
; PRIOR FILING DATE: 2000-10-25
; NUMBER OF SEQ ID NOS: 4
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 3
; LENGTH: 8622
; TYPE: DNA
; ORGANISM: Human
; US-09-867-570-3

```

Query Match	42.1%;	Score 1101;	DB 11;	Length 8622;
Best Local Similarity	100.0%;	Pred. No. 0;		
Matches 1101;	Conservative 0;	Mismatches 0;	Indels 0;	Gaps 0;
Qy	466	GGGTCATCAGACTGCGGGGTTCTGAGCMTGAAATTCACCATCCGAGTCTGGGATCAGAAC	525	
Db	7502	GGGTCATCAGACTGCGGGGTTCTGAGCMTGAAATTCACCATCCGAGTCTGGGATCAGAAC	7561	
Qy	526	TGACACCAATCAAGGAGCGTAGGAGACTCCTTGCTACAGACAGACCCCTGAGCTTACAGG	585	
Db	7562	TGACACCAATCAAGGAGCGTAGGAGACTCCTTGCTACAGACAGACCCCTGAGCTTACAGG	7621	
Qy	586	GGCTGACGTGCATCGTTCCCTTGTGCGGCTGACAGAAAGCGGGTTGCTTGGCTGCC	645	
Db	7622	GGCTGACGTGCATCGTTCCCTTGTGCGGCTGACAGAAAGCGGGTTGCTTGGCTGCC	7681	
Qy	646	TGGGCTGCGGAGTCCGACGAGAACGCTGTCTTCATCTACATCCTCAACCTGGTCCGCGCG	705	
Db	7682	TGGGCTGCGGAGTCCGACGAGAACGCTGTCTTCATCTACATCCTCAACCTGGTCCGCGCG	7741	
Qy	706	ACTTCTCTTCTTAGCGGACCATTAATATGTTGCGCGTTACGCGCTCATCAATATCGCGC	765	
Db	7742	ACTTCTCTTCTTAGCGGACCATTAATATGTTGCGCGTTACGCGCTCATCAATATCGCGC	7801	
Qy	766	ATCCCATCTCCAAATCTCTAGTCTGTGATGACCTTTCCTACTTTATAGGCTTAAGCA	825	
Db	7802	ATCCCATCTCCAAATCTCTAGTCTGTGATGACCTTTCCTACTTTATAGGCTTAAGCA	7861	
Qy	826	TGCTGAGCGCATCAGACCGAGCGCTGCTGTCTCATCTGTGCGCATCTGGTACACT	885	

Db	7862	TGCTGACGGCATATAGACACGAGGCGTGGCTGTCCATCTGTGGCCCATCTGTGACACT	7921
Oy	886	GCCGCCGCCCCAGATACCTGTCAATCGGTCAATGTGTGTCTGTGAGCCCTGTGCCCTGC	945
Db	7922	GCCGCCGCCCCAGATACCTGTCAATCGGTCAATGTGTGTCTGTGAGCCCTGTGCCCTGC	7981
Oy	946	TGCGGAGTATCTTGAGATGATGTTCTGTGACCTTCCTGTTTAGTGATGTCGATTCCTGTTT	1005
Db	7982	TGCGGAGTATCTTGAGATGATGTTCTGTGACCTTCCTGTTTAGTGATGTCGATTCCTGTTT	8041
Oy	1006	GGTGTGAACGTCAGATTTCAATTACATTCGGGTGGCTGTTTTTTATGATGTGTTCTCT	1065
Db	8042	GGTGTGAACGTCAGATTTCAATTACATTCGGGTGGCTGTTTTTTATGATGTGTTCTCT	8101
Oy	1066	GTGGGTCACAGCTGTGTCCTGTGTGACGATTCCTGTGTGATCCCGGAAGATGCGCTGA	1125
Db	8102	GTGGGTCACAGCTGTGTCCTGTGTGACGATTCCTGTGTGATCCCGGAAGATGCGCTGA	8161
Oy	1126	CCAGGCTGTATAGTACCATCTCTCTCAAGTGTGATCTTCTCTCTGTGTGGCTGTCCCT	1185
Db	8162	CCAGGCTGTATAGTACCATCTCTCTCAAGTGTGATCTTCTCTCTGTGTGGCTGTCCCT	8221
Oy	1186	TTTGGCATTCAGTGGGCCCTGTTTTCCAGGATCCACTGTGATTTGGAAGTCTTATTTTGTCT	1245
Db	8222	TTTGGCATTCAGTGGGCCCTGTTTTCCAGGATCCACTGTGATTTGGAAGTCTTATTTTGTCT	8281
Oy	1246	ATGTGCATCTAGTTTCCATTTTCCGTGCGGCTCTTAACAGACGTGCCAACCCCATCATTTT	1305
Db	8282	ATGTGCATCTAGTTTCCATTTTCCGTGCGGCTCTTAACAGACGTGCCAACCCCATCATTTT	8341
Oy	1306	ACTTCTTCGTGGGCTCCTTTAGGACAGGCTCAAAATAGGACGAACCTGAAGCTGTCTCC	1365
Db	8342	ACTTCTTCGTGGGCTCCTTTAGGACAGGCTCAAAATAGGACGAACCTGAAGCTGTCTCC	8401
Oy	1366	AGAGGGCTCTGCAGACACGCTGTGAGGTGATGGAAGGTGAGAGGTGCTTCTCAGAAAA	1425
Db	8402	AGAGGGCTCTGCAGACACGCTGTGAGGTGATGGAAGGTGAGAGGTGCTTCTCAGAAAA	8461
Oy	1426	CCCTGTGAGCTGTGGGGAAGCAGATTTGAGACATGTAGAGAAACCTCTGTCTGTCAACA	1485
Db	8462	CCCTGTGAGCTGTGGGGAAGCAGATTTGAGACATGTAGAGAAACCTCTGTCTGTCAACA	8521
Oy	1486	GGAATTTTGAAGCAATGCTGCCCTGCACACCTTGACAAATTATATGCAATTTTCTTAGCCT	1545
Db	8522	GGAATTTTGAAGCAATGCTGCCCTGCACACCTTGACAAATTATATGCAATTTTCTTAGCCT	8581
Oy	1546	TCTGCTCAGAAATGTCTCAG 1566	
Db	8582	TCTGCTCAGAAATGTCTCAG 8602	

```

RESULT 3
US-10-292-798-1273
: Sequence 1273, Application US/10292798
: Publication No. US20030235833A1
: GENERAL INFORMATION:
: APPLICANT: SUMA, MAKIKO
: APPLICANT: ASAI, KIYOSHI
: APPLICANT: AKIYAMA, YUTAKA
: APPLICANT: ABEIRATANI, HIROYUKI
: TITLE OF INVENTION: GUANOSINE TRIPHOSPHATE-BINDING PROTEIN COUPLED RECEPTORS
: FILE REFERENCE: 084335/166
: CURRENT APPLICATION NUMBER: US/10/292,798
: CURRENT FILING DATE: 2002-11-13
: PRIOR APPLICATION NUMBER: 10/017,161
: PRIOR FILING DATE: 2001-12-18
: PRIOR APPLICATION NUMBER: JP 2001-246789
: PRIOR FILING DATE: 2001-06-18
: NUMBER OF SEQ ID NOS: 2070
: SOFTWARE: patentIn Ver. 2.1
: SEQ ID NO 1273
:
: LENGTH: 1369

```

```

; TYPE: DNA
; ORGANISM: Homo sapiens
; FEATURE:
; LOCATION: source
; FEATURE:
; LOCATION: (1) .. (1369)
; FEATURE:
; NAME/KEY: CDS
; LOCATION: (201) .. (1169)
US-10-292-798-1273

Query Match      41.7%; Score 1093; DB 15; Length 1369;
Best Local Similarity 99.8%; Pred. No. 0;
Matches 1193; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

QY 466 GGGTCATCAGACGTGGGGTTTCTGAGCATGATTCACATCCAGCTTGGGTAACAGAC 525
DB 175 GGGTCATCAGACGTGGGGTTTCTGAGCATGATTCACATCCAGCTTGGGTAACAGAC 234
QY 526 TGACACCAATCAACGGAAGTGAAGAGACTCTTGTCTACAGAGAACCTTGAAGCTTCAAG 585
DB 235 TGACACCAATCAACGGAAGTGAAGAGACTCTTGTCTACAGAGAACCTTGAAGCTTCAAG 294
QY 586 GGGTCATCAGACGTGGGGTTTCTGAGCATGATTCACATCCAGCTTGGGTAACAGAC 645
DB 295 GGGTCATCAGACGTGGGGTTTCTGAGCATGATTCACATCCAGCTTGGGTAACAGAC 354
QY 646 TGGGTCGCGAGATGGGAGAAAGCGTGTCTCATCTACATCTTCAACCTGGTCCGAGCGG 705
DB 355 TGGGTCGCGAGATGGGAGAAAGCGTGTCTCATCTACATCTTCAACCTGGTCCGAGCGG 414
QY 706 ACTTCTCTCTCTTACGCGGCAATTAATGTTGCGCGTTAGCGCTCATCAATATCCGCG 765
DB 415 ACTTCTCTCTCTTACGCGGCAATTAATGTTGCGCGTTAGCGCTCATCAATATCCGCG 474
QY 766 ATCCCATCTCCAAATCTCTCACTCTGTGATGACCTTTCCCTTACTTAAAGCCTTAAGA 825
DB 475 ATCCCATCTCCAAATCTCTCACTCTGTGATGACCTTTCCCTTACTTAAAGCCTTAAGA 534
QY 826 TGTGAGGCGCATGACGACCGGAGCGCTGCTGCTCAATCTGTGAGCGCATCTGTAACACT 885
DB 535 TGTGAGGCGCATGACGACCGGAGCGCTGCTGCTCAATCTGTGAGCGCATCTGTAACACT 594
QY 886 GCCCGCGCGGAGATACCTGTGATGATGATGATGATGATGATGATGATGATGATGATG 945
DB 595 GCCCGCGCGGAGATACCTGTGATGATGATGATGATGATGATGATGATGATGATGATG 654
QY 946 TGGGAGATATCTGAGATGAGATGATGATGATGATGATGATGATGATGATGATGATG 1005
DB 655 TGGGAGATATCTGAGATGAGATGATGATGATGATGATGATGATGATGATGATGATG 714
QY 1006 GGTGGAACAGTCAGATTTCAATTAACATGCGGAGCGTGTGTTTTTATGATGATGATG 1065
DB 715 GGTGGAACAGTCAGATTTCAATTAACATGCGGAGCGTGTGTTTTTATGATGATGATG 774
QY 1066 GTGGGTCCAGGCTGTGCTGTGCTGTGCTGTGCTGTGCTGTGCTGTGCTGTGCTGTG 1125
DB 775 GTGGGTCCAGGCTGTGCTGTGCTGTGCTGTGCTGTGCTGTGCTGTGCTGTGCTGTG 834
QY 1126 CCAAGCTGTGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATG 1185
DB 835 CCAAGCTGTGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATG 894
QY 1186 TTGGCATTCAGGAGCGGCTGTTTTTCAAGATCAGCTGTGATGGAAGCTTAATTTTGTG 1245
DB 895 TTGGCATTCAGGAGCGGCTGTTTTTCAAGATCAGCTGTGATGGAAGCTTAATTTTGTG 954
QY 1246 ATGTGATCTAGATTTCAATTTCTGTGCTGTGCTGTGCTGTGCTGTGCTGTGCTGTG 1305
DB 955 ATGTGATCTAGATTTCAATTTCTGTGCTGTGCTGTGCTGTGCTGTGCTGTGCTGTG 1014
QY 1306 ACTTCTCTGAGGCTCTTTAGGAGCGTCAAAATAGGAGAACTGAAAGCTGTGCTTCC 1365
DB 1306 ACTTCTCTGAGGCTCTTTAGGAGCGTCAAAATAGGAGAACTGAAAGCTGTGCTTCC 1365
```

```

DB 1015 ACTTCTCTGAGGCTCTTTAGGAGCGTCAAAATAGGAGAACTGAAAGCTGTGCTTCC 1074
QY 1366 AGAGGCTCTGAGAGACAGCGCTGAGTGAATGAAGTGAAGGTTGCTTCTCAGAGAA 1425
DB 1075 AGAGGCTCTGAGAGACAGCGCTGAGTGAATGAAGTGAAGGTTGCTTCTCAGAGAA 1134
QY 1426 CCTGAGCTGTGAGAGAGAGATTTGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 1485
DB 1135 CCTGAGCTGTGAGAGAGAGATTTGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 1194
QY 1486 GAGCTTTAGAGAGAGATGCTGAGCGCTGAGCGCTGAGCAATTAATGATTTTCTAGCT 1545
DB 1195 GAGCTTTAGAGAGAGATGCTGAGCGCTGAGCGCTGAGCAATTAATGATTTTCTAGCT 1254
QY 1546 TCTGCTCAGAAATGCTCAGTGTGCTCCTCAAGGCTCTGAAATGATTTTAACT 1605
DB 1255 TCTGCTCAGAAATGCTCAGTGTGCTCCTCAAGGCTCTGAAATGATTTTAACT 1314
QY 1606 GACAGTTGAGTTTCAACCATGGAAGCATTAATGCTGACATCAATGTTTGA 1660
DB 1315 GACAGTTGAGTTTCAACCATGGAAGCATTAATGCTGACATCAATGTTTGA 1369

RESULT 4
US-10-183-116-30
; Sequence 30, Application US/10183116
; Publication No. US2003092035A1
; GENERAL INFORMATION:
; APPLICANT: Anderson, David J.
; APPLICANT: Dong, Xinhong
; APPLICANT: Zylka, Mark
; APPLICANT: Han, Sang-kyou
; TITLE OF INVENTION: PAIN SIGNALING MOLECULES
; FILE REFERENCE: CALTR 4C1CPI
; CURRENT APPLICATION NUMBER: US/10/183,116
; CURRENT FILING DATE: 2002-06-26
; PRIOR APPLICATION NUMBER: US 60/222,344
; PRIOR FILING DATE: 2000-08-01
; PRIOR APPLICATION NUMBER: US 60/202,027
; PRIOR FILING DATE: 2000-05-04
; PRIOR APPLICATION NUMBER: US 09/704,707
; PRIOR FILING DATE: 2000-11-03
; PRIOR APPLICATION NUMBER: US 60/285,493
; PRIOR FILING DATE: 2001-04-19
; PRIOR APPLICATION NUMBER: US 09/849,869
; PRIOR FILING DATE: 2001-05-04
; NUMBER OF SEQ ID NOS: 109
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 30
; LENGTH: 1400
; TYPE: DNA
; ORGANISM: Homo sapiens
; FEATURE:
; NAME/KEY: CDS
; LOCATION: (332) ... (1297)
US-10-183-116-30

Query Match      41.6%; Score 1088; DB 14; Length 1400;
Best Local Similarity 100.0%; Pred. No. 0;
Matches 1088; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 473 CAGACTGGGGTTTCTGAGCATGATTCACATCCAGCTTGGGTAACAGAC 532
DB 313 CAGACTGGGGTTTCTGAGCATGATTCACATCCAGCTTGGGTAACAGAC 372
QY 533 AATCAAGGAGCGTGAAGAGACTCTTGTCTACAGAGAACCTTGAAGCTTCAAGGAGCTGAC 592
DB 373 AATCAAGGAGCGTGAAGAGACTCTTGTCTACAGAGAACCTTGAAGCTTCAAGGAGCTGAC 432
QY 593 GTGATCGTTTCCCTTGTGCGGCTGACAGAGAAAGCGGTTGTGCTGTGCGGCTG 652
DB 433 GTGATCGTTTCCCTTGTGCGGCTGACAGAGAAAGCGGTTGTGCTGTGCGGCTG 492
```

QY 653 CCGCATGCGAGGAAACGCTGCTCCATCTACATCTCTCAACCTGCTCGCGGCACTTCT 712
DB 493 CCGCATGCGAGGAAACGCTGCTCCATCTACATCTCTCAACCTGCTCGCGGCACTTCT 552
QY 713 CTTCCTTAGGCGGCACTTATATATGTCGCGGTACGCGCTCAATATCCGCATCTCCAT 772
DB 553 CTTCCTTAGGCGGCACTTATATATGTCGCGGTACGCGCTCAATATCCGCATCTCCAT 612
QY 773 CTTCCTTAGGCGGCACTTATATGTCGCGGTACGCGCTCAATATCCGCATCTCCAT 832
DB 613 CTTCCTTAGGCGGCACTTATATGTCGCGGTACGCGCTCAATATCCGCATCTCCAT 672
QY 833 CCGCATGCGAGGAAACGCTGCTCCATCTCTCAACCTGCTCGCGGCACTTCT 892
DB 673 CCGCATGCGAGGAAACGCTGCTCCATCTCTCAACCTGCTCGCGGCACTTCT 732
QY 893 CCGCATGCGAGGAAACGCTGCTCCATCTCTCAACCTGCTCGCGGCACTTCT 952
DB 733 CCGCATGCGAGGAAACGCTGCTCCATCTCTCAACCTGCTCGCGGCACTTCT 792
QY 953 TATCTGAGTGAATGTTCTGTCATCTCTGTCATCTCTGTCATCTCTGTCATCTCTGTCAT 1012
DB 793 TATCTGAGTGAATGTTCTGTCATCTCTGTCATCTCTGTCATCTCTGTCATCTCTGTCAT 852
QY 1013 AACGTGAGTGAATGTTCTGTCATCTCTGTCATCTCTGTCATCTCTGTCATCTCTGTCAT 1072
DB 853 AACGTGAGTGAATGTTCTGTCATCTCTGTCATCTCTGTCATCTCTGTCATCTCTGTCAT 912
QY 1073 CAGCTGAGTGAATGTTCTGTCATCTCTGTCATCTCTGTCATCTCTGTCATCTCTGTCAT 1132
DB 913 CAGCTGAGTGAATGTTCTGTCATCTCTGTCATCTCTGTCATCTCTGTCATCTCTGTCAT 972
QY 1133 GTACGTGAGTGAATGTTCTGTCATCTCTGTCATCTCTGTCATCTCTGTCATCTCTGTCAT 1192
DB 973 GTACGTGAGTGAATGTTCTGTCATCTCTGTCATCTCTGTCATCTCTGTCATCTCTGTCAT 1032
QY 1193 TCGAGTGAATGTTCTGTCATCTCTGTCATCTCTGTCATCTCTGTCATCTCTGTCAT 1252
DB 1033 TCGAGTGAATGTTCTGTCATCTCTGTCATCTCTGTCATCTCTGTCATCTCTGTCAT 1092
QY 1253 TCGAGTGAATGTTCTGTCATCTCTGTCATCTCTGTCATCTCTGTCATCTCTGTCAT 1312
DB 1093 TCGAGTGAATGTTCTGTCATCTCTGTCATCTCTGTCATCTCTGTCATCTCTGTCAT 1152
QY 1313 CCGCATGCGAGGAAACGCTGCTCCATCTCTCAACCTGCTCGCGGCACTTCT 1372
DB 1153 CCGCATGCGAGGAAACGCTGCTCCATCTCTCAACCTGCTCGCGGCACTTCT 1212
QY 1373 TCGAGTGAATGTTCTGTCATCTCTGTCATCTCTGTCATCTCTGTCATCTCTGTCAT 1432
DB 1213 TCGAGTGAATGTTCTGTCATCTCTGTCATCTCTGTCATCTCTGTCATCTCTGTCAT 1272
QY 1433 GCTGCGGAGGAAACGCTGCTCCATCTCTCAACCTGCTCGCGGCACTTCT 1492
DB 1273 GCTGCGGAGGAAACGCTGCTCCATCTCTCAACCTGCTCGCGGCACTTCT 1332
QY 1493 GAGAGCATGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 1552
DB 1333 GAGAGCATGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 1392
QY 1553 CAGAAATG 1560
DB 1393 CAGAAATG 1400

RESULT 5
US-10-225-567A-673

; Sequence 673, Application US/10225567A
; Publication No. US20030113798A1
; GENERAL INFORMATION:
; APPLICANT: Lifespan Biosciences
; APPLICANT: Brown, Joseph P.

APPLICANT: Burner, Glenna C.
APPLICANT: Roush, Christine L.
TITLE OF INVENTION: ANTIGENIC PEPTIDES AND ANTIBODIES FOR G PROTEIN-COUPLED RECEPTORS
FILE REFERENCE: 1920-4-4
CURRENT APPLICATION NUMBER: US/10/225,567A
PRIOR FILING DATE: 2001-12-19
PRIOR APPLICATION NUMBER: 60/257,144
NUMBER OF SEQ ID NOS: 2292
SOFTWARE: PatentIn version 3.1
SEQ ID NO 673
LENGTH: 1400
TYPE: DNA
ORGANISM: Homo sapiens
US-10-225-567A-673

Query Match 41.6%; Score 1088; DB 15; Length 1400;
Best Local Similarity 100.0%; Pred. No. 0;
Matches 1088; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 473 CAGCTGAGTGAATGTTCTGTCATCTCTGTCATCTCTGTCATCTCTGTCATCTCTGTCAT 532
DB 313 CAGCTGAGTGAATGTTCTGTCATCTCTGTCATCTCTGTCATCTCTGTCATCTCTGTCAT 372
QY 533 AATCAACGAGTGAAGAGATCTCTGTCATCTCTGTCATCTCTGTCATCTCTGTCAT 592
DB 373 AATCAACGAGTGAAGAGATCTCTGTCATCTCTGTCATCTCTGTCATCTCTGTCAT 432
QY 593 GTGATGCTGTTCCCTGTCATCTCTGTCATCTCTGTCATCTCTGTCATCTCTGTCAT 652
DB 433 GTGATGCTGTTCCCTGTCATCTCTGTCATCTCTGTCATCTCTGTCATCTCTGTCAT 492
QY 653 CCGCATGCGAGGAAACGCTGCTCCATCTCTCAACCTGCTCGCGGCACTTCT 712
DB 493 CCGCATGCGAGGAAACGCTGCTCCATCTCTCAACCTGCTCGCGGCACTTCT 552
QY 713 CTTCCTTAGGCGGCACTTATATATGTCGCGGTACGCGCTCAATATCCGCATCTCCAT 772
DB 553 CTTCCTTAGGCGGCACTTATATATGTCGCGGTACGCGCTCAATATCCGCATCTCCAT 612
QY 773 CTTCCTTAGGCGGCACTTATATATGTCGCGGTACGCGCTCAATATCCGCATCTCCAT 832
DB 613 CTTCCTTAGGCGGCACTTATATATGTCGCGGTACGCGCTCAATATCCGCATCTCCAT 672
QY 833 CCGCATGCGAGGAAACGCTGCTCCATCTCTCAACCTGCTCGCGGCACTTCT 892
DB 673 CCGCATGCGAGGAAACGCTGCTCCATCTCTCAACCTGCTCGCGGCACTTCT 732
QY 893 CCGCATGCGAGGAAACGCTGCTCCATCTCTCAACCTGCTCGCGGCACTTCT 952
DB 733 CCGCATGCGAGGAAACGCTGCTCCATCTCTCAACCTGCTCGCGGCACTTCT 792
QY 953 TATCTGAGTGAATGTTCTGTCATCTCTGTCATCTCTGTCATCTCTGTCATCTCTGTCAT 1012
DB 793 TATCTGAGTGAATGTTCTGTCATCTCTGTCATCTCTGTCATCTCTGTCATCTCTGTCAT 852
QY 1013 AACGTGAGTGAATGTTCTGTCATCTCTGTCATCTCTGTCATCTCTGTCATCTCTGTCAT 1072
DB 853 AACGTGAGTGAATGTTCTGTCATCTCTGTCATCTCTGTCATCTCTGTCATCTCTGTCAT 912
QY 1073 CAGCTGAGTGAATGTTCTGTCATCTCTGTCATCTCTGTCATCTCTGTCATCTCTGTCAT 1132
DB 913 CAGCTGAGTGAATGTTCTGTCATCTCTGTCATCTCTGTCATCTCTGTCATCTCTGTCAT 972
QY 1133 GTACGTGAGTGAATGTTCTGTCATCTCTGTCATCTCTGTCATCTCTGTCATCTCTGTCAT 1192
DB 973 GTACGTGAGTGAATGTTCTGTCATCTCTGTCATCTCTGTCATCTCTGTCATCTCTGTCAT 1032
QY 1193 TCGAGTGAATGTTCTGTCATCTCTGTCATCTCTGTCATCTCTGTCATCTCTGTCAT 1252
DB 1033 TCGAGTGAATGTTCTGTCATCTCTGTCATCTCTGTCATCTCTGTCATCTCTGTCAT 1092
QY 1253 TCGAGTGAATGTTCTGTCATCTCTGTCATCTCTGTCATCTCTGTCATCTCTGTCAT 1312

Db 1093 TCTAGTTCCATTTTCCCTGCTCGCTCTTAACAGCAGTGCACCCCATCATTTACTTCTT 1152
QY 1313 CGTGGGCTCCTTTGGGACGCGTCAAAATAGGAGAAACCTGAAGCTGTCTCCAGAGGAC 1372
Db 1153 CGTGGGCTCCTTTAGGACGCGTCAAAATAGGAGAAACCTGAAGCTGTCTCCAGAGGAC 1212
QY 1373 TCTCAGAGACACGCTGAGTGTAGTAAAGTGAAGGCTGTCTCTCAGAAACCTGTGA 1432
Db 1213 TCTCAGAGACACGCTGAGTGTAGTAAAGTGAAGGCTGTCTCTCAGAAACCTGTGA 1272
QY 1433 GCTGTCCGGAAGCAGATTGAGCAGTGAAGAAACCTTGCCCTGTCAAGACAGACTTT 1492
Db 1273 GCTGTCCGGAAGCAGATTGAGCAGTGAAGAAACCTTGCCCTGTCAAGACAGACTTT 1332
QY 1493 GAGAGCAATGCTGCGCCCTGCACCCCTTGACATTAATGATATTTCTTGTAGCTTGCT 1552
Db 1333 GAGAGCAATGCTGCGCCCTGCACCCCTTGACATTAATGATATTTCTTGTAGCTTGCT 1392
QY 1553 CAGAATG 1560
Db 1393 CAGAATG 1400

RESULT 6
US-09-995-225-19
; Sequence 19, Application US/09995225
; Publication No. US20020193584A1
; GENERAL INFORMATION:
; APPLICANT: Chen, Ruoping
; APPLICANT: Chu, Zhi Liang
; APPLICANT: Dang, Huong T.
; APPLICANT: Lowitz, Kevin P.
; APPLICANT: Pride, Cameron
; TITLE OR INVENTION: Endogenous And No. US20020193584A1-Endogenous Versions of Human C
; FILE REFERENCE: AREN-0308
; CURRENT APPLICATION NUMBER: US/09/995,225
; PRIOR FILING DATE: 2001-11-26
; PRIOR APPLICATION NUMBER: 09/170,496
; PRIOR FILING DATE: 1998-10-13
; PRIOR APPLICATION NUMBER: PCT/US99/23938
; PRIOR FILING DATE: 1998-10-13
; PRIOR APPLICATION NUMBER: 60/253,404
; PRIOR FILING DATE: 2000-11-27
; PRIOR APPLICATION NUMBER: 60/255,366
; PRIOR FILING DATE: 2000-12-12
; PRIOR APPLICATION NUMBER: 60/270,286
; PRIOR FILING DATE: 2001-02-20
; PRIOR APPLICATION NUMBER: 60/282,365
; PRIOR FILING DATE: 2001-04-06
; PRIOR APPLICATION NUMBER: 60/270,266
; PRIOR FILING DATE: 2001-02-20
; PRIOR APPLICATION NUMBER: 60/282,032
; PRIOR FILING DATE: 2001-04-06
; PRIOR APPLICATION NUMBER: 60/282,358
; PRIOR FILING DATE: 2001-04-06
; PRIOR APPLICATION NUMBER: 60/282,356
; PRIOR FILING DATE: 2001-04-06
; PRIOR APPLICATION NUMBER: 60/290,917
; PRIOR FILING DATE: 2001-05-14
; PRIOR APPLICATION NUMBER: 60/309,208
; PRIOR FILING DATE: 2001-07-31
; NUMBER OF SEQ ID NOS: 67
; SOFTWARE: PatentIn version 3.1
; SEQ ID NO 19
; LENGTH: 969
; TYPE: DNA
; ORGANISM: Artificial Sequence
; FEATURE:
; OTHER INFORMATION: No. US20020193584A1el Sequence
US-09-995-225-19

Query Match 37.0%; Score 969; DB 9; Length 969;
Best Local Similarity 100.0%; Pred. No. 0;
Matches 969; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 432 ATGATTTCAACATCCGCTTGGGTACAGAACTGACACCAATCAACGAGCGAGAG 551
Db 1 ATGATTTCAACATCCGCTTGGGTACAGAACTGACACCAATCAACGAGCGAGAG 60
QY 552 ACTCTTGCTCAACAGACAGACCTGAGCTTCAAGGAGGCTGAGTGCATGTTTCCCTTGC 611
Db 61 ACTCTTGCTCAACAGACAGACCTGAGCTTCAAGGAGGCTGAGTGCATGTTTCCCTTGC 120
QY 612 GCGCTGACAGAAACGCGGTTGTGCTTCTGAGCTCTGAGGCTGCGGCAATGCGAGAAAGCT 671
Db 121 GCGCTGACAGAAACGCGGTTGTGCTTCTGAGGCTCTGAGGCTGCGGCAATGCGAGAAAGCT 180
QY 672 GTCTCATCTACATCTCAACCTGGTGGCGGCGAGCTTCTCTTGAAGGCGACATT 731
Db 181 GTCTCATCTACATCTCAACCTGGTGGCGGCGAGCTTCTCTTGAAGGCGACATT 240
QY 732 ATATGTTGCGGTTACGCGCTCATCAATATCCGCAATCCATCTGCAAAATCTGAGTCT 791
Db 241 ATATGTTGCGGTTACGCGCTCATCAATATCCGCAATCCATCTGCAAAATCTGAGTCT 300
QY 792 GTGATGACCTTTCCTACTTTATAGGCTTAAGCATGCTGAGCGGCAATCAAGACGAGCGC 851
Db 301 GTGATGACCTTTCCTACTTTATAGGCTTAAGCATGCTGAGCGGCAATCAAGACGAGCGC 360
QY 852 TGCCGTGCATCCGTTGAGCCCATGCTGACACACTGCGCGCGCCAGATACCTGTCATCG 911
Db 361 TGCCGTGCATCCGTTGAGCCCATGCTGACACACTGCGCGCGCCAGATACCTGTCATCG 420
QY 912 GTCATGTGTGCTGCTCTGAGGCGCTGTCCCTGCTGCGAGATATCTGAGTGTGATTC 971
Db 421 GTCATGTGTGCTGCTCTGAGGCGCTGTCCCTGCTGCGAGATATCTGAGTGTGATTC 480
QY 972 TGTGACTTCTGTTTATGAGTGTGCTGATTTCTGTTGTGTAAGCTGATTTATTACA 1031
Db 481 TGTGACTTCTGTTTATGAGTGTGCTGATTTCTGTTGTGTAAGCTGATTTATTACA 540
QY 1032 ATCGGAGGCTGTTTTTTATGATGAGTGTCTGAGGAGTCAAGCTGATCTGCTGATC 1091
Db 541 ATCGGAGGCTGTTTTTTATGATGAGTGTCTGAGGAGTCAAGCTGATCTGCTGATC 600
QY 1092 AGGATTTCTGTGTGATCCCGAAGATGCGCTGACAGGCTGTACGTGACCATCTCTC 1151
Db 601 AGGATTTCTGTGTGATCCCGAAGATGCGCTGACAGGCTGTACGTGACCATCTCTC 660
QY 1152 ACAGTGCTGCTTCTCTCTGAGGCGCTGCTGAGCTTGTGCAATTCAGTGAGCCCTGTTCC 1211
Db 661 ACAGTGCTGCTTCTCTCTGAGGCGCTGCTGAGCTTGTGCAATTCAGTGAGCCCTGTTCC 720
QY 1212 AGGATCAACCTGAGTTGAAAGTCTTATTTGTATGATGATCTAGTTTCCATTTTCTG 1271
Db 721 AGGATCAACCTGAGTTGAAAGTCTTATTTGTATGATGATCTAGTTTCCATTTTCTG 780
QY 1272 TCCGCTTTTAAACAGAGTGCACACCCCATCTTACTTCTGAGGCTCTCTTAAAGCAG 1331
Db 781 TCCGCTTTTAAACAGAGTGCACACCCCATCTTACTTCTGAGGCTCTCTTAAAGCAG 840
QY 1332 CGTCAAAATAGGACAGACCTGAGTGTCTTCCAGAGGCTGTGACAGACCGCTGAG 1391
Db 841 CGTCAAAATAGGACAGACCTGAGTGTCTTCCAGAGGCTGTGACAGACCGCTGAG 900
QY 1392 GTGATGAAGTGAAGGCTGCTTCTCAGGAAACCTGAGGCTGTGAGGAGAGAGATTG 1451
Db 901 GTGATGAAGTGAAGGCTGCTTCTCAGGAAACCTGAGGCTGTGAGGAGAGAGATTG 960
QY 1452 GAGCAGTGA 1460
Db 961 GAGCAGTGA 969

RESULT 7
US-09-995-225-19
; Sequence 19, Application US/09995225
; Publication No. US20030139588A9
; GENERAL INFORMATION:
; APPLICANT: Chen, Ruoping
; APPLICANT: Chu, Zhi Liang
; APPLICANT: Dang, Huong T.
; APPLICANT: Lowitz, Kevin P.
; APPLICANT: Pridle, Cameron
; TITLE OF INVENTION: Endogenous And No. US20030139588A9-Endogenous Versions of Human G
; TITLE OF INVENTION: Receptors
; FILE REFERENCE: AREN-0308
; CURRENT FILING DATE: 2001-11-26
; PRIOR APPLICATION NUMBER: US/09/995,225
; PRIOR APPLICATION NUMBER: 09/170,496
; PRIOR FILING DATE: 1998-10-13
; PRIOR APPLICATION NUMBER: PCT/US99/23938
; PRIOR FILING DATE: 1998-10-13
; PRIOR APPLICATION NUMBER: 60/253,404
; PRIOR FILING DATE: 2000-11-27
; PRIOR APPLICATION NUMBER: 60/255,366
; PRIOR FILING DATE: 2000-12-12
; PRIOR APPLICATION NUMBER: 60/270,286
; PRIOR FILING DATE: 2001-02-20
; PRIOR APPLICATION NUMBER: 60/282,365
; PRIOR FILING DATE: 2001-04-06
; PRIOR APPLICATION NUMBER: 60/270,266
; PRIOR FILING DATE: 2001-02-20
; PRIOR APPLICATION NUMBER: 60/282,032
; PRIOR FILING DATE: 2001-04-06
; PRIOR APPLICATION NUMBER: 60/282,358
; PRIOR FILING DATE: 2001-04-06
; PRIOR APPLICATION NUMBER: 60/282,356
; PRIOR FILING DATE: 2001-04-06
; PRIOR APPLICATION NUMBER: 60/290,917
; PRIOR FILING DATE: 2001-05-14
; PRIOR APPLICATION NUMBER: 60/309,208
; PRIOR FILING DATE: 2001-07-31
; NUMBER OF SEQ ID NOS: 67
; SOFTWARE: PatentIn version 3.1
; SEQ ID NO 19
; LENGTH: 969
; TYPE: DNA
; ORGANISM: Artificial Sequence
; FEATURE:
; OTHER INFORMATION: No. US20030139588A9el Sequence
US-09-995-225-19

Query Match 37.0%; Score 969; DB 10; Length 969;
Best Local Similarity 100.0%; Pred. No. 0;
Matches 969; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 492 ATGATTCACACATCCAGCTTGGGTACAGAACTGACACCAATCAACGAGCTGAGAG 551
DB 1 ATGATTCACACATCCAGCTTGGGTACAGAACTGACACCAATCAACGAGCTGAGAG 60
QY 552 ACTGCTTGTACAGACGAGCTTGTGCTGAGCTTCAAGGAGCTGAGCTGAGCTTGTGCTG 611
DB 61 ACTGCTTGTACAGACGAGCTTGTGCTGAGCTTCAAGGAGCTGAGCTGAGCTTGTGCTG 120
QY 612 GCGGTGACAGAAACGGGTTGTGCTGAGCTTCTGGGCTGCGGAGCTGAGGAGAGCT 671
DB 121 GCGGTGACAGAAACGGGTTGTGCTGAGCTTCTGGGCTGCGGAGCTGAGGAGAGCT 180
QY 672 GTCTCATCTACATCTCAACCTGTGTGCGGCGGAGCTTCTTCTTCTTCTTCTTCTTCTTCT 731
DB 181 GTCTCATCTACATCTCAACCTGTGTGCGGCGGAGCTTCTTCTTCTTCTTCTTCTTCTTCT 240
QY 732 ATATGTTGCGGCTTACGCTCATCAATATCGGCAATCCCATCTTCAAAATCTTCAATCTTCT 791
DB 241 ATATGTTGCGGCTTACGCTCATCAATATCGGCAATCCCATCTTCAAAATCTTCAATCTTCT 300

QY 792 GTGATGACCTTTCCTTACTTTATAGCCTTAGACAGCTGAGCGGCATCAGACCGAGCGC 851
DB 301 GTGATGACCTTTCCTTACTTTATAGCCTTAGACAGCTGAGCGGCATCAGACCGAGCGC 360
QY 852 TGCCTGTGATCTCTGTGAGCCCATCTGTGATACATGCGCGCGCCCAATATCTGTATCG 911
DB 361 TGCCTGTGATCTCTGTGAGCCCATCTGTGATACATGCGCGCGCCCAATATCTGTATCG 420
QY 912 GTCATGTGTGCTGCTGAGGAGCTGAGGAGCTGAGGAGCTGAGGAGCTGAGGAGCTGAG 971
DB 421 GTCATGTGTGCTGCTGAGGAGCTGAGGAGCTGAGGAGCTGAGGAGCTGAGGAGCTGAG 480
QY 972 TGTGATCTCTGTTTATGAGTGTGATCTGTTTGTGATGAAAGCTGATTCATTACA 1031
DB 481 TGTGATCTCTGTTTATGAGTGTGATCTGTTTGTGATGAAAGCTGATTCATTACA 540
QY 1032 ATGCGTGTGCTGTTTATGAGTGTGATCTGTTTGTGATGAAAGCTGATTCATTACA 1091
DB 541 ATGCGTGTGCTGTTTATGAGTGTGATCTGTTTGTGATGAAAGCTGATTCATTACA 600
QY 1092 AGGATCTCTGTGATCTGAGGAGCTGAGGAGCTGAGGAGCTGAGGAGCTGAGGAGCTG 1151
DB 601 AGGATCTCTGTGATCTGAGGAGCTGAGGAGCTGAGGAGCTGAGGAGCTGAGGAGCTG 660
QY 1152 ACAGTGCTGTGCTCTCTCTCTGAGCTGAGCTGAGCTGAGCTGAGCTGAGCTGAGCTG 1211
DB 661 ACAGTGCTGTGCTCTCTCTCTGAGCTGAGCTGAGCTGAGCTGAGCTGAGCTGAGCTG 720
QY 1212 AGGATCACTGTGATGAAAGCTTATTTTGTGATGATGATGATGATGATGATGATGATG 1271
DB 721 AGGATCACTGTGATGAAAGCTTATTTTGTGATGATGATGATGATGATGATGATGATG 780
QY 1272 TCGGCTTTTAAACGAGATGCAACCCCATATTACTTCTGAGGAGCTCTTTAGGAGAG 1331
DB 781 TCGGCTTTTAAACGAGATGCAACCCCATATTACTTCTGAGGAGCTCTTTAGGAGAG 840
QY 1332 CGTCAAAATAGGAGGAGAACTGAAAGCTGTTCTCCAGAGGAGCTGAGGAGGAGCTGAG 1391
DB 841 CGTCAAAATAGGAGGAGAACTGAAAGCTGTTCTCCAGAGGAGCTGAGGAGGAGCTGAG 900
QY 1392 GTGATGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 1451
DB 901 GTGATGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 960
QY 1452 GAGCAGTGA 1460
DB 961 GAGCAGTGA 969

RESULT 8
US-10-401-397A-1
; Sequence 1, Application US/10401397A
; Publication No. US20030212001A1
; GENERAL INFORMATION:
; APPLICANT: Peri, Krishna G.
; APPLICANT: Moffett, Serge
; APPLICANT: Abran, Daniel
; TITLE OF INVENTION: METHODS AND COMPOUNDS FOR PREVENTION AND TREATMENT OF ELEVATED
; FILE REFERENCE: 4518/1M674US1
; CURRENT FILING DATE: 2003-03-27
; PRIOR APPLICATION NUMBER: US/10/401,397A
; PRIOR FILING DATE: 2002-03-27
; NUMBER OF SEQ ID NOS: 8
; SOFTWARE: PatentIn version 3.1
; SEQ ID NO 1
; LENGTH: 969
; TYPE: DNA
; ORGANISM: Homo sapiens
; FEATURE:
; NAME/KEY: CDS
; LOCATION: (1)..(969)

OTHER INFORMATION: coding sequence for polypeptide
US-10-401-397A-1

Query Match 35.0%; Score 916; DB 15; Length 969;
Best Local Similarity 99.9%; Pred. No. 0;
Matches 966; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

```
QY 492 ATGATTCACCATCCAGTCTTGGGTACAGAACTGACCAATCAACGAGCGTAGAG 551
DB 1 ATGATTCACCATCCAGTCTTGGGTACAGAACTGACCAATCAACGAGCGTAGAG 60
QY 552 ACTCTTCTTACACAGACCTTGAAGCTTCAAGGGCTGACGTGATCTCTTCCCTTGC 611
DB 61 ACTCTTCTTACACAGACCTTGAAGCTTCAAGGGCTGACGTGATCTCTTCCCTTGC 120
QY 612 GCGCTGACAGAAAGCGGGTGTGCTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 671
DB 121 GCGCTGACAGAAAGCGGGTGTGCTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 180
QY 672 GTCTCATCTACATCTCAACCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 731
DB 181 GTCTCATCTACATCTCAACCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 240
QY 732 ATATGTTGCGGTACGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 791
DB 241 ATATGTTGCGGTACGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 300
QY 792 GTGATGACCTTCCCTTCTTATAGGCTTAAAGCATGCTGAGCGCATCAACGAGCGC 851
DB 301 GTGATGACCTTCCCTTCTTATAGGCTTAAAGCATGCTGAGCGCATCAACGAGCGC 360
QY 852 TGCCGTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 911
DB 361 TGCCGTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 420
QY 912 GTGATGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 971
DB 421 GTGATGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 480
QY 972 TGTGATCTCTGTTAGTGTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 1031
DB 481 TGTGATCTCTGTTAGTGTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 540
QY 1032 ATCCGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 1091
DB 541 ATCCGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 600
QY 1092 AGGATTCCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 1151
DB 601 AGGATTCCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 660
QY 1152 ACAAGTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 1211
DB 661 ACAAGTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 720
QY 1212 AGGATTCACCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 1271
DB 721 AGGATTCACCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 780
QY 1272 TCCGCTCTTAAAGAGTGGCAACCCCATCTTCTGCTGCTGCTGCTGCTGCTGCT 1331
DB 781 TCCGCTCTTAAAGAGTGGCAACCCCATCTTCTGCTGCTGCTGCTGCTGCTGCT 840
QY 1332 CGTCAAAATAGGCAACCTGAGTGTCTTCAAGAGGCTCTGAGAGCAAGCTGAG 1391
DB 841 CGTCAAAATAGGCAACCTGAGTGTCTTCAAGAGGCTCTGAGAGCAAGCTGAG 900
QY 1392 GTGATGAGAGTGAAGGCTGCTTCTTCAAGAAACCTGAGAGCTGCGGAAAGCATTTG 1451
DB 901 GTGATGAGAGTGAAGGCTGCTTCTTCAAGAAACCTGAGAGCTGCGGAAAGCATTTG 960
QY 1452 GAGCAGT 1458
DB |||||
```

DB 961 GAGCAGT 967

```
RESULT 9
US-10-391-074-1
; Sequence 1, Application US/10391074
; Publication No. US20040038345A1
; GENERAL INFORMATION:
; APPLICANT: Nehls, Michael
; APPLICANT: Matleier, Frank
; TITLE OF INVENTION: No. US20040038345A1e1 Human Seven-Transmembrane Receptors
; FILE REFERENCE: 7205-0008-00-000
; CURRENT APPLICATION NUMBER: US/10/391,074
; NUMBER OF SEQ ID NOS: 4
; SOFTWARE: FastSeq for Windows Version 3.0
; SEQ ID NO 1
; LENGTH: 969
; TYPE: DNA
; ORGANISM: Homo sapien
US-10-391-074-1
```

Query Match 33.1%; Score 867; DB 16; Length 969;

Best Local Similarity 99.9%; Pred. No. 0;
Matches 917; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

```
QY 492 ATGATTCACCATCCAGTCTTGGGTACAGAACTGACCAATCAACGAGCGTAGAG 551
DB 1 ATGATTCACCATCCAGTCTTGGGTACAGAACTGACCAATCAACGAGCGTAGAG 60
QY 552 ACTCTTCTTACACAGACCTTGAAGCTTCAAGGGCTGACGTGATCTCTTCCCTTGC 611
DB 61 ACTCTTCTTACACAGACCTTGAAGCTTCAAGGGCTGACGTGATCTCTTCCCTTGC 120
QY 612 GCGCTGACAGAAAGCGGGTGTGCTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 671
DB 121 GCGCTGACAGAAAGCGGGTGTGCTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 180
QY 672 GTCTCATCTACATCTCAACCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 731
DB 181 GTCTCATCTACATCTCAACCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 240
QY 732 ATATGTTGCGGTACGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 791
DB 241 ATATGTTGCGGTACGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 300
QY 792 GTGATGACCTTCCCTTCTTATAGGCTTAAAGCATGCTGAGCGCATCAACGAGCGC 851
DB 301 GTGATGACCTTCCCTTCTTATAGGCTTAAAGCATGCTGAGCGCATCAACGAGCGC 360
QY 852 TGCCGTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 911
DB 361 TGCCGTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 420
QY 912 GTGATGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 971
DB 421 GTGATGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 480
QY 972 TGTGATCTCTGTTAGTGTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 1031
DB 481 TGTGATCTCTGTTAGTGTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 540
QY 1032 ATCCGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 1091
DB 541 ATCCGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 600
QY 1092 AGGATTCCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 1151
DB 601 AGGATTCCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 660
QY 1152 ACAAGTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 1211
DB 661 ACAAGTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 720
```

QY 1212 AGATCCACCTGATGTAAGTCTTATTTTGTGATGATCTAGTTTCCATTTTCTG 1271
DB 721 AGATCCACCTGATGTAAGTCTTATTTTGTGATGATCTAGTTTCCATTTTCTG 780
QY 1272 TCCGCTCTTAAACAGAGTGGCAACCCCATCTTACTTCTTGTGGGCTCTTTAGGAG 1331
DB 781 TCCGCTCTTAAACAGAGTGGCAACCCCATCTTACTTCTTGTGGGCTCTTTAGGAG 840
QY 1332 CGTCAAAATAGGAGAACTGAGCTGTTCTCGAGAGGCTCGAGAGAGGCTGAG 1391
DB 841 CGTCAAAATAGGAGAACTGAGCTGTTCTCGAGAGGCTCGAGAGAGGCTGAG 900
QY 1392 GTGATGAGAGTGGAGG 1409
DB 901 GTGATGAGAGTGGAGG 918
RESULT 10
US-10-219-834-7
; Sequence 7, Application US/10219834
; Publication No. US20030096751A1
; GENERAL INFORMATION:
; APPLICANT: Bristol-Myers Squibb Company
; TITLE OF INVENTION: G-PROTEIN COUPLED RECEPTOR POLYNUCLEOTIDES AND METHODS OF USE THE
; FILE REFERENCE: D0191 NP
; CURRENT APPLICATION NUMBER: US/10/219, 834
; CURRENT FILING DATE: 2002-08-15, 834
; PRIOR APPLICATION NUMBER: US 60/313, 658
; PRIOR FILING DATE: 2001-08-20
; PRIOR APPLICATION NUMBER: US 60/340, 703
; PRIOR FILING DATE: 2001-10-30
; PRIOR APPLICATION NUMBER: US 60/318, 675
; PRIOR FILING DATE: 2001-09-12
; PRIOR APPLICATION NUMBER: US 60/355, 596
; PRIOR FILING DATE: 2002-02-06
; PRIOR APPLICATION NUMBER: US 60/333, 417
; PRIOR FILING DATE: 2001-11-26
; PRIOR APPLICATION NUMBER: US 60/338, 367
; PRIOR FILING DATE: 2001-12-06
; NUMBER OF SEQ ID NOS: 192
; SOFTWARE: PatentIn version 3.1
; SEQ ID NO 7
; LENGTH: 1997
; TYPE: DNA
; ORGANISM: Homo sapiens
US-10-219-834-7
Query Match 32.7%; Score 856; DB 14; Length 1997;
Best Local Similarity 100.0%; Pred. No. 0;
Matches 856; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 1387 CTGAGGTGATGTAAGTGGAGGTGCTTCTCTCAGAAACCTGAGCTGTGGAGCA 1446
DB 1082 CTGAGGTGATGTAAGTGGAGGTGCTTCTCTCAGAAACCTGAGCTGTGGAGCA 1441
QY 1447 GATTGAGCAGTGAAGAACTCTGCTGCTGAGCAGAGACTTTAGAGCAATGCTGC 1506
DB 1142 GATTGAGCAGTGAAGAACTCTGCTGCTGAGCAGAGACTTTAGAGCAATGCTGC 1201
QY 1507 CTTGCGACCTTGAATTTATGATTTTCTTAGGCTTGTGCTGAGAAATGCTCAG 1566
DB 1202 CTTGCGACCTTGAATTTATGATTTTCTTAGGCTTGTGCTGAGAAATGCTCAG 1261
QY 1567 TGGTCCCTCAAGCTCTTGAATGATTTTATCTTAACCTGACAGTTTCACTTCA 1626
DB 1262 TGGTCCCTCAAGCTCTTGAATGATTTTATCTTAACCTGACAGTTTCACTTCA 1321
QY 1627 TGGAAAGCATTAAGTCTGACAGTACATGTTTGGATTCTCTGATATTTACATAT 1686
DB 1322 TGGAAAGCATTAAGTCTGACAGTACATGTTTGGATTCTCTGATATTTACATAT 1381
QY 1687 TCCCTGTATCTTGAAGTGAATCTTCTACTGAACATTTTCTGACTTTTCAATGTA 1746

DB 1382 TCCCTGTATCTTGCACGTGAATCTTCTCTACTGAACATTTTCTGACTTTTCAATGTA 1441
QY 1747 ATAAAGAGTGGTGGCCCAACCCCAAAACCTCTTATATCTTGTCTTCACTGAT 1806
DB 1442 ATAAAGAGTGGTGGCCCAACCCCAAAACCTCTTATATCTTGTCTTCACTGAT 1501
QY 1807 AGTATCAAAAGAGAGATTCCTTATTAATCTGACATATGTTCCCTGAAATCATGT 1866
DB 1502 AGTATCAAAAGAGAGATTCCTTATTAATCTGACATATGTTCCCTGAAATCATGT 1561
QY 1867 TCCCTTTATGACTGAGGCAATTAATGAGTGAAGCTCAATCTTAAATGAGGTT 1926
DB 1562 TCCCTTTATGACTGAGGCAATTAATGAGTGAAGCTCAATCTTAAATGAGGTT 1621
QY 1927 TGTACCTCTTAATTCATTAATGATTAATCTGACATTAAGCAAAATTAAGACTTGA 1986
DB 1622 TGTACCTCTTAATTCATTAATGATTAATCTGACATTAAGCAAAATTAAGACTTGA 1681
QY 1987 GATTCTCCCTTCAATTAATTAAGAGTCTTGAATGTTTATGAATAGCCCTCTGTA 2046
DB 1682 GATTCTCCCTTCAATTAATTAAGAGTCTTGAATGTTTATGAATAGCCCTCTGTA 1741
QY 2047 TTTGTCCAGACAGTGTGACATGTTGGCTTGTCTTACTGAAGACATGTTGCCCT 2106
DB 1742 TTTGTCCAGACAGTGTGACATGTTGGCTTGTCTTACTGAAGACATGTTGCCCT 1801
QY 2107 TCCCTTGAAGACGTGTAATGTTTATTAATGCTTCTCTGACATTAATGAGTGA 2166
DB 1802 TCCCTTGAAGACGTGTAATGTTTATTAATGCTTCTCTGACATTAATGAGTGA 1861
QY 2167 GCTATTAATATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 2226
DB 1862 GCTATTAATATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 1921
QY 2227 AAGTGAATATATCTT 2242
DB 1922 AAGTGAATATATCTT 1937
RESULT 11
US-10-237-467-9
; Sequence 9, Application US/10237467
; Publication No. US20030186324A1
; GENERAL INFORMATION:
; APPLICANT: Liao, Jiayu
; APPLICANT: Gray, Nathanael S.
; APPLICANT: Caldwell, Jeremy C.
; APPLICANT: Schultz, Peter G.
; APPLICANT: IBM LLC
; TITLE OF INVENTION: Sensory Neuron Receptors
; FILE REFERENCE: 021288-001300US
; CURRENT APPLICATION NUMBER: US/10/237,467
; CURRENT FILING DATE: 2003-01-14
; PRIOR APPLICATION NUMBER: US 60/317, 879
; PRIOR FILING DATE: 2001-09-07
; NUMBER OF SEQ ID NOS: 18
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 9
; LENGTH: 909
; TYPE: DNA
; ORGANISM: Homo sapiens
; FEATURES:
; OTHER INFORMATION: dorsal root ganglia G-protein coupled receptor (GPCR)
; OTHER INFORMATION: 7 (DRG7) (NT009307)
US-10-237-467-9
Query Match 28.5%; Score 747; DB 15; Length 909;
Best Local Similarity 99.8%; Pred. No. 0;
Matches 847; Conservative 0; Mismatches 2; Indels 0; Gaps 0;
QY 612 GCGCTGACAGAAAGCGGTTGTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 671


```
Db 61 GCGCTGACGAGAACCGGCTGTGTCTCTGAGCTCTGAGGCTGCCGACATGCGACGAAAGCT 120
Qy 672 GTCTCATCTACATCTCAACCTGATGCGCGCGAGCTTCTCTTACGCGGCAACT 731
Db 121 GTCTCATCTACATCTCAACCTGATGCGCGCGAGCTTCTCTTACGCGGCAACT 180
Qy 732 ATATGTTGCGCGTTAGCGCTCATCAATATCCGCAATCCCATCTCCAAATCTCAGTCT 791
Db 181 ATATGTTGCGCGTTAGCGCTCATCAATATCCGCAATCCCATCTCCAAATCTCAGTCT 240
Qy 792 GTGATGACCTTCCCTACTTTATAGAGCCCAAGCATGATGAGGCGCATGACGCGAGGCG 851
Db 241 GTGATGACCTTCCCTACTTTATAGAGCCCAAGCATGATGAGGCGCATGACGCGAGGCG 300
Qy 852 TGCCCTGATCATCTGTGAGCCCATCTGTATACATGCGCGCGCGCCAGATACCTGTATCG 911
Db 301 TGCCCTGATCATCTGTGAGCCCATCTGTATACATGCGCGCGCGCCAGATACCTGTATCA 360
Qy 912 GTCATGTGTCTGTGCTGTGAGGCTGTCTGCTGCTGCGGAGTATCTGTGAGTGTTC 971
Db 361 GTCATGTGTCTGTGCTGTGAGGCTGTCTGCTGCTGCGGAGTATCTGTGAGTGTTC 420
Qy 972 TGTGATCTTCTGTGTATGATGATGATGATGATGATGATGATGATGATGATGATGAT 1031
Db 421 TGTGATCTTCTGTGTATGATGATGATGATGATGATGATGATGATGATGATGATGAT 480
Qy 1032 ATCCGATGCTGTGTATGATGATGATGATGATGATGATGATGATGATGATGATGAT 1091
Db 481 ATCCGATGCTGTGTATGATGATGATGATGATGATGATGATGATGATGATGATGAT 540
Qy 1092 AGAATCTCTGTGTATGATGATGATGATGATGATGATGATGATGATGATGATGAT 1151
Db 541 AGAATCTCTGTGTATGATGATGATGATGATGATGATGATGATGATGATGATGAT 600
Qy 1152 ACAATGCTGTGTATGATGATGATGATGATGATGATGATGATGATGATGATGAT 1211
Db 601 ACAATGCTGTGTATGATGATGATGATGATGATGATGATGATGATGATGATGAT 660
Qy 1212 AGAATCCATGATGATGATGATGATGATGATGATGATGATGATGATGAT 1271
Db 661 AGAATCCATGATGATGATGATGATGATGATGATGATGATGATGATGAT 720
Qy 1272 TCCGCTCTTAAAGAGTGTGCAACCCCATCTTAACTTCTGTGAGCTCTTTAGGCG 1331
Db 721 TCCGCTCTTAAAGAGTGTGCAACCCCATCTTAACTTCTGTGAGCTCTTTAGGCG 780
Qy 1332 CGTCAAAATAGGCAAGACTGATGATGATGATGATGATGATGATGATGATGATGAT 1391
Db 781 CGTCAAAATAGGCAAGACTGATGATGATGATGATGATGATGATGATGATGATGAT 840
Qy 1392 GTGATGAAAGTGTGAGGCTGTCTGATGAAACCTGTGAGCTGTGCGGAAAGCAGATTG 1451
Db 841 GTGATGAAAGTGTGAGGCTGTCTGATGAAACCTGTGAGCTGTGCGGAAAGCAGATTG 900
Qy 1452 GAGCAGTGA 1460
Db 901 GAGCAGTGA 909

RESULT 12
US-10-101-510-239
; Sequence 239, Application US/10101510
; Publication No. US20030148295A1
; GENERAL INFORMATION:
; APPLICANT: WAN, JACKSON
; APPLICANT: WANG, YIXIN
; TITLE OF INVENTION: EXPRESSION PROFILES AND METHODS OF USE
; FILE REFERENCE: 15117.0012
; CURRENT APPLICATION NUMBER: US/10/101,510
; PRIORITY FILING DATE: 2002-03-20
; PRIOR APPLICATION NUMBER: 60/276,947
; PRIORITY FILING DATE: 2001-03-20
; NUMBER OF SEQ ID NOS: 805
```

```
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 239
; LENGTH: 769
; TYPE: DNA
; ORGANISM: Homo sapiens
; FEATURE:
; NAME/KEY: modified_base
; LOCATION: (750)
; OTHER INFORMATION: a, t, c, g, other or unknown
US-10-101-510-239

Query Match      26.7%; Score 698; DB 15; Length 769;
Best Local Similarity 99.9%; Pred. No. 0;
Matches 748; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

Qy 762 CGCCATCCCATCTCCAAATCTCAGTCTGTGATGACCTTCCCTATTTATAGGCTTA 821
Db 1 CGCCATCCCATCTCCAAATCTCAGTCTGTGATGACCTTCCCTATTTATAGGCTTA 60
Qy 822 AGCATGCTGAGGCGCATCAGACCGAGGCTGCTTCATCTGTGAGCCCATCTGTATC 881
Db 61 AGCATGCTGAGGCGCATCAGACCGAGGCTGCTTCATCTGTGAGCCCATCTGTATC 120
Qy 882 CACTGCGCGCGCCGAGATACCTGTGATGATGATGATGATGATGATGATGATGATGAT 941
Db 121 CACTGCGCGCGCCGAGATACCTGTGATGATGATGATGATGATGATGATGATGATGAT 180
Qy 942 CTGCTGCGGAGATCTGAGTGTGATGATGATGATGATGATGATGATGATGATGATGAT 1001
Db 181 CTGCTGCGGAGATCTGAGTGTGATGATGATGATGATGATGATGATGATGATGATGAT 240
Qy 1002 GTTGTGTGAAACGTCAGATTTCAATCAATGCGGTGTGTGTGTGTGTGTGTGTGT 1061
Db 241 GTTGTGTGAAACGTCAGATTTCAATCAATGCGGTGTGTGTGTGTGTGTGTGTGT 300
Qy 1062 CTCTGTGAGTCAAGCTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGT 1121
Db 301 CTCTGTGAGTCAAGCTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGT 360
Qy 1122 CTGACGAGCTGTAGTACCATCTCTCAAGTGTGTGTGTGTGTGTGTGTGTGTGTGT 1181
Db 361 CTGACGAGCTGTAGTACCATCTCTCAAGTGTGTGTGTGTGTGTGTGTGTGTGTGT 420
Qy 1182 CCGTTTGGCATTCAGTGTGAGGCTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGT 1241
Db 421 CCGTTTGGCATTCAGTGTGAGGCTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGT 480
Qy 1242 TGTGATGTGATCTAGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGT 1301
Db 481 TGTGATGTGATCTAGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGT 540
Qy 1302 ATTATCTTCTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGT 1361
Db 541 ATTATCTTCTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGT 600
Qy 1362 CTCAGAGGCTCTGTGAGACAGCGCTGTGATGATGATGATGATGATGATGATGATGAT 1421
Db 601 CTCAGAGGCTCTGTGAGACAGCGCTGTGATGATGATGATGATGATGATGATGATGAT 660
Qy 1422 GAAACCTGTGAGCTGTGCGGAAAGCAGATTGAGAGCAGTGAAGAAACCTGTGCTGCA 1481
Db 661 GAAACCTGTGAGCTGTGCGGAAAGCAGATTGAGAGCAGTGAAGAAACCTGTGCTGCA 720
Qy 1482 GACAGACTTGTGAGGCAATGCTGCGCTG 1510
Db 721 GACAGACTTGTGAGGCAATGCTGCGCTG 749

RESULT 13
US-10-017-161-1599
; Sequence 1599, Application US/10017161
; Publication No. US2003014368A1
; GENERAL INFORMATION:
```


APPLICANT: SUMA, MAKIKO
APPLICANT: ASAI, KIYOSHI
APPLICANT: AKIYAMA, YUTAKA
APPLICANT: ABEURATANI, HIROYUKI
TITLE OF INVENTION: NOVEL G PROTEIN-COUPLED RECEPTORS
FILE REFERENCE: 084335/0152
CURRENT APPLICATION NUMBER: US/10/017,161
CURRENT FILING DATE: 2002-12-18
PRIOR APPLICATION NUMBER: JP 2001/246789
PRIOR FILING DATE: 2001-06-18
NUMBER OF SEQ ID NOS: 2430
SOFTWARE: PatentIn Ver. 2.1
SEQ ID NO 1599
LENGTH: 1370
TYPE: DNA
ORGANISM: Homo sapiens
FEATURE:
NAME/KEY: source
LOCATION: (1)..(1370)
FEATURE:
NAME/KEY: CDS
LOCATION: (201)..(619)
FEATURE:
NAME/KEY: CDS
LOCATION: (996)..(1170)
US-10-017-161-1599

Query Match 16.6%; Score 434; DB 15; Length 1370;
Best Local Similarity 100.0%; Pred. No. 1.2e-210;
Matches 434; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1227 TGAAGCTTATTTGTCATGTCATCTAGTTTCCATTTCCGTCGCTTAAGC 1286
DB 937 TGAAGCTTATTTGTCATGTCATCTAGTTTCCATTTCCGTCGCTTAAGC 996
QY 1287 AGTGCCAAACCCATCATTTCTTCTGCGCTCTTTAGGACGCTCAAAATAGCAG 1346
DB 997 AGTGCCAAACCCATCATTTCTTCTGCGCTCTTTAGGACGCTCAAAATAGCAG 1096
QY 1347 AACCTGAAGCTGTTCTCCAGAGGCTCTGAGACACGCTGAGTGATGAAGTGA 1406
DB 1057 AACCTGAAGCTGTTCTCCAGAGGCTCTGAGACACGCTGAGTGATGAAGTGA 1116
QY 1407 GGGTGGCTTCTCTGAGAAACCTGAGAGCTGCGGGAAGAGATTTGAGAGAGAGA 1466
DB 1117 GGGTGGCTTCTCTGAGAAACCTGAGAGCTGCGGGAAGAGATTTGAGAGAGAGA 1176
QY 1467 ACCTGCGCTGTCAGACAGGACTTTGAGAGCAATGCTGCGCTGCAACCTTGAACAATTA 1526
DB 1177 ACCTGCGCTGTCAGACAGGACTTTGAGAGCAATGCTGCGCTGCAACCTTGAACAATTA 1236
QY 1527 TATGCAATTTTCTTACGCTTCTGCTGCTGCAAAATGTCAGTGTCTCTCAAGTCTTGA 1586
DB 1237 TATGCAATTTTCTTACGCTTCTGCTGCTGCAAAATGTCAGTGTCTCTCAAGTCTTGA 1296
QY 1587 ATGAGTGTATCTAACCTGACAGTTCACCTGCAAGTTCACCTGCAAGTTCACCTGCA 1646
DB 1297 ATGAGTGTATCTAACCTGACAGTTCACCTGCAAGTTCACCTGCAAGTTCACCTGCA 1356
QY 1647 GTACAAATGTTTGA 1660
DB 1357 GTACAAATGTTTGA 1370

RESULT 14
US-10-313-542-164
Sequence 164, Application US/10313542
Publication No. US20030120057A1
GENERAL INFORMATION:
APPLICANT: Roopa, Reddy
APPLICANT: Guegler, Karl, J.
APPLICANT: Au-Young, Janice
TITLE OF INVENTION: COMPOSITION FOR DETECTION OF GENES ENCODING MEMBRANE-ASSOCIATED F

FILE REFERENCE: PA-0013 US
CURRENT APPLICATION NUMBER: US/10/313,542
CURRENT FILING DATE: 2002-12-05
PRIOR APPLICATION NUMBER: US/09/495,050
PRIOR FILING DATE: 2000-01-31
PRIOR APPLICATION NUMBER: 60/118,318
PRIOR FILING DATE: 1999-02-01
NUMBER OF SEQ ID NOS: 305
SOFTWARE: PERL Program
SEQ ID NO 164
LENGTH: 291
TYPE: DNA
ORGANISM: Homo sapiens
FEATURE:
NAME/KEY: misc feature
OTHER INFORMATION: Incyte ID No. US20030120057A1 1909132CT1
US-10-313-542-164

Query Match 11.1%; Score 291; DB 15; Length 291;
Best Local Similarity 100.0%; Pred. No. 1.5e-137;
Matches 291; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 762 CGCCATCCATCTCCAAATCTCAGTCTGTGATGACCTTTCCCTATATAGGCTTA 821
DB 1 CGCCATCCATCTCCAAATCTCAGTCTGTGATGACCTTTCCCTATATAGGCTTA 60
QY 822 AGCATGCTGAGCCCATCAGACCGAGCGCTGCTGTCATCTGAGGCGCATCTGTGATC 881
DB 61 AGCATGCTGAGCCCATCAGACCGAGCGCTGCTGTCATCTGAGGCGCATCTGTGATC 120
QY 882 CACTGCGCGCGCCAGATACCTGTCATCGATGATGTCCTGCTGAGGCGCTGTC 941
DB 121 CACTGCGCGCGCCAGATACCTGTCATCGATGATGTCCTGCTGAGGCGCTGTC 180
QY 942 CTGCTGCGAGTATCTGAGTGAATGTCGTGACTTCTGTTAGTGTGATTTCT 1001
DB 181 CTGCTGCGAGTATCTGAGTGAATGTCGTGACTTCTGTTAGTGTGATTTCT 240
QY 1002 GTTGTGTCGAACGTCAGATTCATTACAATGCGCGGCTGATTTTCTTA 1052
DB 241 GTTGTGTCGAACGTCAGATTCATTACAATGCGCGGCTGATTTTCTTA 291

RESULT 15
US-10-305-720-330
Sequence 330, Application US/10305720
Publication No. US20040010136A1
GENERAL INFORMATION:
APPLICANT: Au-Young, Janice K.; Seilhamer, Jeffrey J.
TITLE OF INVENTION: Composition for the Detection of Signaling Pathway Gene Expression
FILE REFERENCE: PA-0002-1 CON
CURRENT APPLICATION NUMBER: US/10/305,720
CURRENT FILING DATE: 2002-11-26
PRIOR APPLICATION NUMBER: 09/016,434
PRIOR FILING DATE: 1998-01-30
NUMBER OF SEQ ID NOS: 1490
SOFTWARE: PERL Program
SEQ ID NO 330
LENGTH: 275
TYPE: DNA
ORGANISM: Homo sapiens
FEATURE:
NAME/KEY: misc feature
OTHER INFORMATION: Incyte ID No. US20040010136A1 1909132
US-10-305-720-330

Query Match 10.5%; Score 275; DB 16; Length 275;
Best Local Similarity 100.0%; Pred. No. 2.3e-129;
Matches 275; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 762 CGCCATCCATCTCCAAATCTCAGTCTGTGATGACCTTTCCCTATATAGGCTTA 821
DB 1 CGCCATCCATCTCCAAATCTCAGTCTGTGATGACCTTTCCCTATATAGGCTTA 60

GenCore version 5.1.6
Copyright (c) 1993 - 2004 CompuGen Ltd.

OM nucleic - nucleic search, using sw model

Run on: October 29, 2004, 19:13:03 ; Search time 9581 Seconds
(without alignments)
11080.481 Million cell updates/sec

Title: US-09-867-570-1

Perfect score: 2618
Sequence: 1 aacaatgcgcgcgaattcgg.....gaaaaaaaaaaaaaaaaa 2618

Scoring table: OLIGO_NUC
Gapop 60.0 , Gapext 60.0

Searched: 44308572 seqs, 20275418765 residues

Word size : 20

Total number of hits satisfying chosen parameters: 189184

Minimum DB seq length: 0

Maximum DB seq length: 200000000

Post-processing: Listing first 45 summaries

Database :

```
Pending Patents NA Main: *
1: /cgn2_6/prodata/1/pna/PCRU1 COMB.seq: *
2: /cgn2_6/prodata/1/pna/PCRU2 COMB.seq: *
3: /cgn2_6/prodata/1/pna/US06 COMB.seq: *
4: /cgn2_6/prodata/1/pna/US07 COMB.seq: *
5: /cgn2_6/prodata/1/pna/US08 COMB.seq: *
6: /cgn2_6/prodata/1/pna/US081 COMB.seq: *
7: /cgn2_6/prodata/1/pna/US082 COMB.seq: *
8: /cgn2_6/prodata/1/pna/US083 COMB.seq: *
9: /cgn2_6/prodata/1/pna/US084 COMB.seq: *
10: /cgn2_6/prodata/1/pna/US085 COMB.seq: *
11: /cgn2_6/prodata/1/pna/US086 COMB.seq: *
12: /cgn2_6/prodata/1/pna/US087 COMB.seq: *
13: /cgn2_6/prodata/1/pna/US088 COMB.seq: *
14: /cgn2_6/prodata/1/pna/US089 COMB.seq: *
15: /cgn2_6/prodata/1/pna/US090 COMB.seq: *
16: /cgn2_6/prodata/1/pna/US091 COMB.seq: *
17: /cgn2_6/prodata/1/pna/US092A COMB.seq: *
18: /cgn2_6/prodata/1/pna/US092B COMB.seq: *
19: /cgn2_6/prodata/1/pna/US093A COMB.seq: *
20: /cgn2_6/prodata/1/pna/US093B COMB.seq: *
21: /cgn2_6/prodata/1/pna/US094 COMB.seq: *
22: /cgn2_6/prodata/1/pna/US095A COMB.seq: *
23: /cgn2_6/prodata/1/pna/US095B COMB.seq: *
24: /cgn2_6/prodata/1/pna/US095C COMB.seq: *
25: /cgn2_6/prodata/1/pna/US095D COMB.seq: *
26: /cgn2_6/prodata/1/pna/US096A COMB.seq: *
27: /cgn2_6/prodata/1/pna/US096B COMB.seq: *
28: /cgn2_6/prodata/1/pna/US096C COMB.seq: *
29: /cgn2_6/prodata/1/pna/US096D COMB.seq: *
30: /cgn2_6/prodata/1/pna/US096E COMB.seq: *
31: /cgn2_6/prodata/1/pna/US097A COMB.seq: *
32: /cgn2_6/prodata/1/pna/US097B COMB.seq: *
33: /cgn2_6/prodata/1/pna/US097C COMB.seq: *
34: /cgn2_6/prodata/1/pna/US098A COMB.seq: *
35: /cgn2_6/prodata/1/pna/US098B COMB.seq: *
36: /cgn2_6/prodata/1/pna/US098C COMB.seq: *
37: /cgn2_6/prodata/1/pna/US098D COMB.seq: *
38: /cgn2_6/prodata/1/pna/US099A COMB.seq: *
39: /cgn2_6/prodata/1/pna/US099B COMB.seq: *
40: /cgn2_6/prodata/1/pna/US099C COMB.seq: *
41: /cgn2_6/prodata/1/pna/US099D COMB.seq: *
42: /cgn2_6/prodata/1/pna/US099E COMB.seq: *
43: /cgn2_6/prodata/1/pna/US099F COMB.seq: *
44: /cgn2_6/prodata/1/pna/US099G COMB.seq: *
45: /cgn2_6/prodata/1/pna/US100A COMB.seq: *
46: /cgn2_6/prodata/1/pna/US100B COMB.seq: *
47: /cgn2_6/prodata/1/pna/US101A COMB.seq: *
48: /cgn2_6/prodata/1/pna/US101B COMB.seq: *
49: /cgn2_6/prodata/1/pna/US102A COMB.seq: *
50: /cgn2_6/prodata/1/pna/US102B COMB.seq: *
51: /cgn2_6/prodata/1/pna/US103A COMB.seq: *
52: /cgn2_6/prodata/1/pna/US103B COMB.seq: *
53: /cgn2_6/prodata/1/pna/US104A COMB.seq: *
54: /cgn2_6/prodata/1/pna/US104B COMB.seq: *
55: /cgn2_6/prodata/1/pna/US105A COMB.seq: *
56: /cgn2_6/prodata/1/pna/US105B COMB.seq: *
57: /cgn2_6/prodata/1/pna/US106A COMB.seq: *
58: /cgn2_6/prodata/1/pna/US107A COMB.seq: *
59: /cgn2_6/prodata/1/pna/US107B COMB.seq: *
60: /cgn2_6/prodata/1/pna/US107C COMB.seq: *
61: /cgn2_6/prodata/1/pna/US107D COMB.seq: *
62: /cgn2_6/prodata/1/pna/US108A COMB.seq: *
63: /cgn2_6/prodata/1/pna/US108B COMB.seq: *
64: /cgn2_6/prodata/1/pna/US109A COMB.seq: *
65: /cgn2_6/prodata/1/pna/US6000 COMB.seq: *
66: /cgn2_6/prodata/1/pna/US6001 COMB.seq: *
67: /cgn2_6/prodata/1/pna/US6002 COMB.seq: *
68: /cgn2_6/prodata/1/pna/US6003 COMB.seq: *
69: /cgn2_6/prodata/1/pna/US6004 COMB.seq: *
70: /cgn2_6/prodata/1/pna/US6005 COMB.seq: *
71: /cgn2_6/prodata/1/pna/US6006 COMB.seq: *
72: /cgn2_6/prodata/1/pna/US6007 COMB.seq: *
73: /cgn2_6/prodata/1/pna/US6008 COMB.seq: *
74: /cgn2_6/prodata/1/pna/US6009 COMB.seq: *
75: /cgn2_6/prodata/1/pna/US6010 COMB.seq: *
76: /cgn2_6/prodata/1/pna/US6011 COMB.seq: *
77: /cgn2_6/prodata/1/pna/US6012 COMB.seq: *
78: /cgn2_6/prodata/1/pna/US6013 COMB.seq: *
79: /cgn2_6/prodata/1/pna/US6015 COMB.seq: *
80: /cgn2_6/prodata/1/pna/US6016 COMB.seq: *
81: /cgn2_6/prodata/1/pna/US6017 COMB.seq: *
82: /cgn2_6/prodata/1/pna/US6018 COMB.seq: *
83: /cgn2_6/prodata/1/pna/US6019 COMB.seq: *
84: /cgn2_6/prodata/1/pna/US6020 COMB.seq: *
85: /cgn2_6/prodata/1/pna/US6021 COMB.seq: *
86: /cgn2_6/prodata/1/pna/US6022 COMB.seq: *
87: /cgn2_6/prodata/1/pna/US6023 COMB.seq: *
88: /cgn2_6/prodata/1/pna/US6023A COMB.seq: *
89: /cgn2_6/prodata/1/pna/US6023B COMB.seq: *
90: /cgn2_6/prodata/1/pna/US6024 COMB.seq: *
91: /cgn2_6/prodata/1/pna/US6025 COMB.seq: *
92: /cgn2_6/prodata/1/pna/US6026 COMB.seq: *
93: /cgn2_6/prodata/1/pna/US6027 COMB.seq: *
94: /cgn2_6/prodata/1/pna/US6028 COMB.seq: *
95: /cgn2_6/prodata/1/pna/US6029 COMB.seq: *
96: /cgn2_6/prodata/1/pna/US6030 COMB.seq: *
97: /cgn2_6/prodata/1/pna/US6031 COMB.seq: *
98: /cgn2_6/prodata/1/pna/US6032 COMB.seq: *
99: /cgn2_6/prodata/1/pna/US6033 COMB.seq: *
100: /cgn2_6/prodata/1/pna/US6033A COMB.seq: *
101: /cgn2_6/prodata/1/pna/US6035 COMB.seq: *
102: /cgn2_6/prodata/1/pna/US6036 COMB.seq: *
103: /cgn2_6/prodata/1/pna/US6037 COMB.seq: *
104: /cgn2_6/prodata/1/pna/US6038 COMB.seq: *
105: /cgn2_6/prodata/1/pna/US6039 COMB.seq: *
106: /cgn2_6/prodata/1/pna/US6040 COMB.seq: *
107: /cgn2_6/prodata/1/pna/US6041 COMB.seq: *
108: /cgn2_6/prodata/1/pna/US6042 COMB.seq: *
109: /cgn2_6/prodata/1/pna/US6043 COMB.seq: *
110: /cgn2_6/prodata/1/pna/US6044 COMB.seq: *
111: /cgn2_6/prodata/1/pna/US6045 COMB.seq: *
112: /cgn2_6/prodata/1/pna/US6046 COMB.seq: *
113: /cgn2_6/prodata/1/pna/US6047 COMB.seq: *
114: /cgn2_6/prodata/1/pna/US6048 COMB.seq: *
115: /cgn2_6/prodata/1/pna/US6049 COMB.seq: *
116: /cgn2_6/prodata/1/pna/US6050 COMB.seq: *
```


Best Local Similarity 100.0%; Pred. No. 0; Mismatches 0; Indels 0; Gaps 0;
Matches 2618; Conservative 0;

1 AACATTGCGGGAATTCGGGACGAGATGAATCTAGTTTAAAGGCTGAGACCT 60
1 AACAATTCGCGGAATTCGGGACGAGATGAATCTAGTTTAAAGGCTGAGACCT 60
61 CCTCCCTCTCTTAACTCTCTCTCTCAACATGTGAGACGCTCGCTCCCTTTGCTTT 120
61 CCTCCCTCTCTTAACTCTCTCTCTCAACATGTGAGACGCTCGCTCCCTTTGCTTT 120
121 CACGAGATTGGAAGCTTCTGAGAGCTTCCCGAAGAGAGAGCTGATGCTTTGTA 180
121 CACGAGATTGGAAGCTTCTGAGAGCTTCCCGAAGAGAGAGCTGATGCTTTGTA 180
181 CAGTCTGAGAGCTTATGAGAGCTTAAACCATTTCTCTATAAATTTCCAGCTCAG 240
181 CAGTCTGAGAGCTTATGAGAGCTTAAACCATTTCTCTATAAATTTCCAGCTCAG 240
241 TATTTCTTTTGAACAATTTGAGAAATGAATTAACACAGACAGAGAGAGATGAA 300
241 TATTTCTTTTGAACAATTTGAGAAATGAATTAACACAGACAGAGAGAGATGAA 300
301 ATCCCAAGTGTCTTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCT 360
301 ATCCCAAGTGTCTTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCT 360
361 TCCACGAGAAACGAGAAATTAAGAAATCCCACTGATGTGATACATAGAACCACTCT 420
361 TCCACGAGAAACGAGAAATTAAGAAATCCCACTGATGTGATACATAGAACCACTCT 420
421 TGGAGTCAAAACAGATTAAGAAATGAAGAAATCTCATGAGTCAATCAAGCTG 480
421 TGGAGTCAAAACAGATTAAGAAATGAAGAAATCTCATGAGTCAATCAAGCTG 480
481 GGTTCGAGAGATGATTCACCATCCAGTCTTGGGTACAGAACTGACACCAATCAAG 540
481 GGTTCGAGAGATGATTCACCATCCAGTCTTGGGTACAGAACTGACACCAATCAAG 540
541 GACGTGAGAGAGCTCTTGTCTACAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 600
541 GACGTGAGAGAGCTCTTGTCTACAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 600
601 TTTCCCTTGTGCGGCTGACAGAGAAACGCGGTTGTGCTGAGCTCTGAGCTGCGGAT 660
601 TTTCCCTTGTGCGGCTGACAGAGAAACGCGGTTGTGCTGAGCTCTGAGCTGCGGAT 660
661 GCAGAGAGCTGTCTCCATCTACATCTCAACCTGTGCGGAGCGGAGCTTCTCTCTTA 720
661 GCAGAGAGCTGTCTCCATCTACATCTCAACCTGTGCGGAGCGGAGCTTCTCTCTTA 720
721 GCGGACCAATTAATAGTTGCGGTTACGCGTCAATCAATTCGCGCATCCCATCTCCAAA 780
721 GCGGACCAATTAATAGTTGCGGTTACGCGTCAATCAATTCGCGCATCCCATCTCCAAA 780
781 TCTCTAGTCTGTGAGAACCTTTCCCTATCTTATAGGCTTAAGCATGCTGAGCGCATCA 840
781 TCTCTAGTCTGTGAGAACCTTTCCCTATCTTATAGGCTTAAGCATGCTGAGCGCATCA 840
841 GCACCGAGCGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 900
841 GCACCGAGCGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 900
901 ACCTGTCAATCGGATCAATGATGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 960
901 ACCTGTCAATCGGATCAATGATGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 960
961 AGTGAATTTCTGTGACTTCTGTTAGTGTGCTGATTTCTGTTGTTGTTGTTGTTG 1020
961 AGTGAATTTCTGTGACTTCTGTTAGTGTGCTGATTTCTGTTGTTGTTGTTGTTG 1020
1021 ATTTCATTACATCGCGTGTGCTGTTTATATGTTGTTCTCTGTGGGTCAAGCTG 1080
1021 ATTTCATTACATCGCGTGTGCTGTTTATATGTTGTTCTCTGTGGGTCAAGCTG 1080

1021 ATTTCATTACATCGCGTGTGCTGTTTATATGTTGTTCTCTGTGGGTCAAGCTG 1080
1081 TCTCTGTGTGAGATTCCTGTGTGATCCCGAAGATCCGCTGACAGAGCTGATGTA 1140
1081 TCTCTGTGTGAGATTCCTGTGTGATCCCGAAGATCCGCTGACAGAGCTGATGTA 1140
1141 CCATCTCTCTCAAGATGCTGCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCT 1200
1141 CCATCTCTCTCAAGATGCTGCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCT 1200
1201 CCCTGTTTCCAGATTCACCTGATGTAAGTCTTATTTGATCATGTGATCAATGTT 1260
1201 CCCTGTTTCCAGATTCACCTGATGTAAGTCTTATTTGATCATGTGATCAATGTT 1260
1261 CCATTTCTCTGCTGCTCTTAAACAGAGTGCACACCCCATCATTTACTTCTGCTG 1320
1261 CCATTTCTCTGCTGCTCTTAAACAGAGTGCACACCCCATCATTTACTTCTGCTG 1320
1321 CCTTTAGGAGAGCTCAAAATAGGAGAACTGAAAGCTGTTCTCAAGAGGCTCTGAG 1380
1321 CCTTTAGGAGAGCTCAAAATAGGAGAACTGAAAGCTGTTCTCAAGAGGCTCTGAG 1380
1381 ACAGGCTGAGGTGATGAAAGGTGAGAGGAGGCTTCTCAGAGAAACCTGAGAGCTG 1440
1381 ACAGGCTGAGGTGATGAAAGGTGAGAGGAGGCTTCTCAGAGAAACCTGAGAGCTG 1440
1441 GAAGAGATTTGAGAGAGTGAAGAGAAACCTCTGCTGCTGACAGAGACTTTGAGAG 1500
1441 GAAGAGATTTGAGAGAGTGAAGAGAAACCTCTGCTGCTGACAGAGACTTTGAGAG 1500
1501 TGCTGCTGCTGACACCTTTGACATTAATGCAATTTTGTAGCTTCTGCTGAGAAATG 1560
1501 TGCTGCTGCTGACACCTTTGACATTAATGCAATTTTGTAGCTTCTGCTGAGAAATG 1560
1561 TCTCAAGTGTCTCAAGAGTCTTCAAGATAGATTTTCTAAGCTGACAGAGTGTG 1620
1561 TCTCAAGTGTCTCAAGAGTCTTCAAGATAGATTTTCTAAGCTGACAGAGTGTG 1620
1621 CACCATGAGAAAGATTAATGCTGACAGATCAATGTTTGAATTCCTTATTAATCAAT 1680
1621 CACCATGAGAAAGATTAATGCTGACAGATCAATGTTTGAATTCCTTATTAATCAAT 1680
1681 AATTTTCCCTGTTATCTTGAAGTCAATCTTCCAGTCAACATTTTCTGCACTTTTC 1740
1681 AATTTTCCCTGTTATCTTGAAGTCAATCTTCCAGTCAACATTTTCTGCACTTTTC 1740
1741 ATTGTATTAAGAGAGTGTGCTGCAACACCTTAAACCTTCTTATTAATCTTCTCTA 1800
1741 ATTGTATTAAGAGAGTGTGCTGCAACACCTTAAACCTTCTTATTAATCTTCTCTA 1800
1801 CCTGATAGTATCAAAAGAGAAATTCCTTATTAATCTGACATGATGTTCCCTGAAA 1860
1801 CCTGATAGTATCAAAAGAGAAATTCCTTATTAATCTGACATGATGTTCCCTGAAA 1860
1861 TCAATGCTCTTTATGACTGAGAGATTAATGCAAGTTGGAAGTCAATTTCTTAATAGT 1920
1861 TCAATGCTCTTTATGACTGAGAGATTAATGCAAGTTGGAAGTCAATTTCTTAATAGT 1920
1921 GAGTTCTGCTACCTGTAATTCATGAAATCTGAGATTAAGAGAAATTAATGACCTTA 1980
1921 GAGTTCTGCTACCTGTAATTCATGAAATCTGAGATTAAGAGAAATTAATGACCTTA 1980
1981 GAGAGAGATTCCTCCCTCATTAATAACAGTCTTAAGAAATGTTTATTAAGATGCTCTC 2040
1981 GAGAGAGATTCCTCCCTCATTAATAACAGTCTTAAGAAATGTTTATTAAGATGCTCTC 2040
2041 CTGTCAATTTGTCACAGATGCTGACATGTTGGCTTGTCTGATTAAGAAATGCTG 2100
2041 CTGTCAATTTGTCACAGATGCTGACATGTTGGCTTGTCTGATTAAGAAATGCTG 2100
2101 GCCCCTTCCCTTGAAGCTGATGATGTTCTTATTAAGCTTCTCTGAGCTTAATGACTAG 2160
2101 GCCCCTTCCCTTGAAGCTGATGATGTTCTTATTAAGCTTCTCTGAGCTTAATGACTAG 2160

[illegible]

```

RESULT 3
US-60-190-928-1
; Sequence 1, Application US/60190928
; GENERAL INFORMATION:
; APPLICANT: Bonazzi, Vivien
; TITLE OF INVENTION: ISOLATED HUMAN G-PROTEIN COUPLED
; TITLE OF INVENTION: RECEPTOR HAVING HOMOLOGY TO THE MAS PROTO-ONCOGENE, NUCLEOTIC
; TITLE OF INVENTION: ACID MOLECULES ENCODING THIS HUMAN GPCR PROTEIN, AND USES
; TITLE OF INVENTION: THEROEF
; FILE REFERENCE: CIL000387
; CURRENT APPLICATION NUMBER: US/60/190,928
; CURRENT FILING DATE: 2000-03-21
; NUMBER OF SEQ ID NOS: 2
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 1
; LENGTH: 2618
; TYPE: DNA
; ORGANISM: HUMAN
US-60-190-928-1

Query Match      100.0%; Score 2618; DB 84; Length 2618;
Best Local Similarity 100.0%; Pred. No. 0;
Matches 2618; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY      1 AACCAATGCCGCCGAATTGGCGACGACGATGGAATCTAGCTGTGTTTAAAGCGGTGACACT 60
DB      1 AACCAATGCCGCCGAATTGGCGACGACGATGGAATCTAGCTGTGTTTAAAGCGGTGACACT 60

QY      CCTCCCTCTCTCTTAACCTCTGCTCTCACACATGTGAGAGCGCTCGCTCCCTCTTGACCTTT 120
DB      61 CCTCCCTCTCTCTTAACCTCTGCTCTCACACATGTGAGAGCGCTCGCTCCCTCTTGACCTTT 120

QY      121 CACCAAGATTGGAAGCTTCTTGAGGCGCTCCCGAAGACGAAGCTGCTATGCTTCTGTA 180
DB      121 CACCAAGATTGGAAGCTTCTTGAGGCGCTCCCGAAGACGAAGCTGCTATGCTTCTGTA 180

QY      181 CAGCTGTGAGAGCTATTAGCCAGTTAAACCCATTTCTTCAATAATTTCCAGTCTCAGG 240
DB      181 CAGCTGTGAGAGCTATTAGCCAGTTAAACCCATTTCTTCAATAATTTCCAGTCTCAGG 240

```

OY	241	TATTTCTTTTGTAGAAATTTGGAAATGAACTPAATCAAGACAGAGAGCAGAGAAATGAAA	300
Db	241	TATTTCTTTTGTAGAAATTTGGAAATGAACTPAATCAAGACAGAGAGCAGAGAAATGAAA	300
OY	301	ATCCCAAGAGTCTTTCCTGCTGTCTTCCAGTCTCTCTGCTGTCTCTCCAGTGTCTCAAT	360
Db	301	ATCCCAAGAGTCTTTCCTGCTGTCTTCCAGTCTCTCTGCTGTCTCTCCAGTGTCTCAAT	360
OY	361	TTCCACCGAAGAACCGAAGAAATTAAGAAATCCCACTGAATGTGTATATGAAGGCCATCTCT	420
Db	361	TTCCACCGAAGAACCGAAGAAATTAAGAAATCCCACTGAATGTGTATATGAAGGCCATCTCT	420
OY	421	TGGATGTCAAAACAGATTAAGAAAGATGGAACCAATCCTCATGGGTCAATCAGACTGCG	480
Db	421	TGGATGTCAAAACAGATTAAGAAAGATGGAACCAATCCTCATGGGTCAATCAGACTGCG	480
OY	481	GGTTTCTGAGCATGGAATTCACATCCACAGCTTGGGTACAGAACTGACACAAATCAACG	540
Db	481	GGTTTCTGAGCATGGAATTCACATCCACAGCTTGGGTACAGAACTGACACAAATCAACG	540
OY	541	GAGGTGAGGAGACTCTTGTCTACAGAGAGACCTGAGCTTCAACGAGGCTGACGTGACATCG	600
Db	541	GAGGTGAGGAGACTCTTGTCTACAGAGAGACCTGAGCTTCAACGAGGCTGACGTGACATCG	600
OY	601	TTTTCCCTTGTGCGGCTGACAGAGAACCGGGTGTGCTCTGGCTCTGAGGCTGCCGATGC	660
Db	601	TTTTCCCTTGTGCGGCTGACAGAGAACCGGGTGTGCTCTGGCTCTGAGGCTGCCGATGC	660
OY	661	GCAGAGAACGCTGTCTCCATCTACATCTCTCAACCTGTGTGCGGCGGACATTCCTCTTCCCTTA	720
Db	661	GCAGAGAACGCTGTCTCCATCTACATCTCTCAACCTGTGTGCGGCGGACATTCCTCTTCCCTTA	720
OY	721	GCGGCCACATTAATATGTTCGGCGGTAAAGCCCTCATCAATATCCGCATCCCATCTCCAAA	780
Db	721	GCGGCCACATTAATATGTTCGGCGGTAAAGCCCTCATCAATATCCGCATCCCATCTCCAAA	780
OY	781	TCCTCAAGTCTGTGATGACCTTTCCCTACCTTATATAGGCTTAAGCATGTGTGAGCGCCATCA	840
Db	781	TCCTCAAGTCTGTGATGACCTTTCCCTACCTTATATAGGCTTAAGCATGTGTGAGCGCCATCA	840
OY	841	GCACCGAGCGGTGCTGTCCATCTGTGTGGCCCATCTGTGTACCATGTGCGCGCGCCCGCAT	900
Db	841	GCACCGAGCGGTGCTGTCCATCTGTGTGGCCCATCTGTGTACCATGTGCGCGCGCCCGCAT	900
OY	901	ACCTGTACATGCGTCATGTGTCTGTGCTCTGTGGCCCTGTCTCCCTGCTGCGAGATATCTGCG	960
Db	901	ACCTGTACATGCGTCATGTGTCTGTGCTCTGTGGCCCTGTCTCCCTGCTGCGAGATATCTGCG	960
OY	961	AGTGGAGATTTCTGTGACTTCTCTGTTAATGTGTGTGTATCTGTGTGTGTGAAACGTACG	1020
Db	961	AGTGGAGATTTCTGTGACTTCTCTGTTAATGTGTGTGTATCTGTGTGTGTGAAACGTACG	1020
OY	1021	ATTTCATTACATGCGGTGCTGTGTTTTTTTATATGTGTCTCTGTGTGCTCAGCCTGCG	1080
Db	1021	ATTTCATTACATGCGGTGCTGTGTTTTTTTATATGTGTCTCTGTGTGCTCAGCCTGCG	1080
OY	1081	TCCTGTGTGTGTCAGGATTTCTGTGTGATCCCGGAAAGATGCGGCTGACCAAGCTGTACGTGA	1140
Db	1081	TCCTGTGTGTGTCAGGATTTCTGTGTGATCCCGGAAAGATGCGGCTGACCAAGCTGTACGTGA	1140
OY	1141	CCATTCCTCTCAGAGTGTGGTCTTCCCTCTGTGTGGCTGCGCTTTTGCCATTCAGTGTGGG	1200
Db	1141	CCATTCCTCTCAGAGTGTGGTCTTCCCTCTGTGTGGCTGCGCTTTTGCCATTCAGTGTGGG	1200
OY	1201	CCCTGTGTTTCCAGGATCACTGTGATGGAAGTCTTAATTTGTGATGTGATCTAGATTT	1260
Db	1201	CCCTGTGTTTCCAGGATCACTGTGATGGAAGTCTTAATTTGTGATGTGATCTAGATTT	1260
OY	1261	CCATTTTCTGTGCTCGGCTTTTAAACGAGTGCACCCCATCATTAATCTTCTGTGTGGGCT	1320
Db	1261	CCATTTTCTGTGCTCGGCTTTTAAACGAGTGCACCCCATCATTAATCTTCTGTGTGGGCT	1320

QY 1321 CCTTGAAGCAGCTCAAAATAGGAGCAAGCTGAGTCTCCAGAGGCTGACAG 1380
DB 1321 CCTTGAAGCAGCTCAAAATAGGAGCAAGCTGAGTCTCCAGAGGCTGACAG 1380
QY 1381 ACAGGCTGAGTGAATGAAGTGGCTTCTCCAGAGAAACCTGAGGCTGACAG 1440
DB 1381 ACAGGCTGAGTGAATGAAGTGGCTTCTCCAGAGAAACCTGAGGCTGACAG 1440
QY 1441 GAAGCAGATGAGAGAGAGAGAAACCTGAGGCTGACAGAGAGCTTGAAGCA 1500
DB 1441 GAAGCAGATGAGAGAGAGAGAAACCTGAGGCTGACAGAGAGCTTGAAGCA 1500
QY 1501 TGCTGCTGCTGACCTTGAAGCAATTAATGCAATTTTCTGAGCTTCTGAGAA 1560
DB 1501 TGCTGCTGCTGACCTTGAAGCAATTAATGCAATTTTCTGAGCTTCTGAGAA 1560
QY 1561 TCTGAGTGGTCCCTCAAGGCTCTGCAATGATGTTATCAACCTGAGAGTGGAC 1620
DB 1561 TCTGAGTGGTCCCTCAAGGCTCTGCAATGATGTTATCAACCTGAGAGTGGAC 1620
QY 1621 CACCATTGAGAAAGCAATTAATGAGAGCAATGTTGAGTCTCTGATATTAACAA 1680
DB 1621 CACCATTGAGAAAGCAATTAATGAGAGCAATGTTGAGTCTCTGATATTAACAA 1680
QY 1681 ACATTTTCTGCTGATATGCTGACATGATTTTCTGAGCACTTTTCTGACCTTTC 1740
DB 1681 ACATTTTCTGCTGATATGCTGACATGATTTTCTGAGCACTTTTCTGACCTTTC 1740
QY 1741 ATTGTAATTAAGAGGAGTGGCTGCAAGACCTTAACCTCTTATATCTGTTTCTTA 1800
DB 1741 ATTGTAATTAAGAGGAGTGGCTGCAAGACCTTAACCTCTTATATCTGTTTCTTA 1800
QY 1801 CCTGATAGTATCAAAAGAGAGATTCCTTATTAATCTGACATATGTTCCCTGAAA 1860
DB 1801 CCTGATAGTATCAAAAGAGAGATTCCTTATTAATCTGACATATGTTCCCTGAAA 1860
QY 1861 TCATGTTCCCTTTATGACTGAGAGCAATGCTGAGTGGAGAGCTCAATTTCAATTA 1920
DB 1861 TCATGTTCCCTTTATGACTGAGAGCAATGCTGAGTGGAGAGCTCAATTTCAATTA 1920
QY 1921 GAGTTCGCTACCTTAATTCATTAATTCAGATTCAGATTAAGCAAAATTAAGCTTA 1980
DB 1921 GAGTTCGCTACCTTAATTCATTAATTCAGATTCAGATTAAGCAAAATTAAGCTTA 1980
QY 1981 GAGAGAGATTCCTCTCATTAATAACAGCTTGAATGTTTGAATAGCCCTCTC 2040
DB 1981 GAGAGAGATTCCTCTCATTAATAACAGCTTGAATGTTTGAATAGCCCTCTC 2040
QY 2041 CTGTCATTTGTCACAGAGAGTGTGACATGTTGCTGCTGCTGCTGCTGCTGCTG 2100
DB 2041 CTGTCATTTGTCACAGAGAGTGTGACATGTTGCTGCTGCTGCTGCTGCTGCTG 2100
QY 2101 GGGCTTCCTCTGAGAGCTGTTGATTTAGCTCTTCCGAGCAATTAAGCAATTA 2160
DB 2101 GGGCTTCCTCTGAGAGCTGTTGATTTAGCTCTTCCGAGCAATTAAGCAATTA 2160
QY 2161 TGAGAGAGCTTAATAATATGTCACAGATTCATTTGAGCAATTAAGCAATTA 2220
DB 2161 TGAGAGAGCTTAATAATATGTCACAGATTCATTTGAGCAATTAAGCAATTA 2220
QY 2221 ATTTTAAGTGAATTAATCTTGAAGAAACATTTATTAATCACTTAAGCAATTA 2280
DB 2221 ATTTTAAGTGAATTAATCTTGAAGAAACATTTATTAATCACTTAAGCAATTA 2280
QY 2281 GTAGAGAGATTCCTCATCTTCAAGGTTTGTATTAATGTTCTGATTAAGCAATTA 2340
DB 2281 GTAGAGAGATTCCTCATCTTCAAGGTTTGTATTAATGTTCTGATTAAGCAATTA 2340
QY 2341 TAGTTTATGCTGTTTATCATAGAGAGCAAACTGAAAATCTGACCTTTGATAGCA 2400
DB 2341 TAGTTTATGCTGTTTATCATAGAGAGCAAACTGAAAATCTGACCTTTGATAGCA 2400
QY 2401 TCTCATTAATGATCTGATTAATACTTAAGTGGTACAGAAATCTGATTAAGCTGT 2460

DB 2401 TCTCATTAATGATCTGATTAATACTTAAGTGGTACAGAAATCTGATTAAGCTGT 2460
QY 2461 GACATACATGACCTGAGAAATTTGCTTAAGAGAAATTAAGCAGACCCCAATATT 2520
DB 2461 GACATACATGACCTGAGAAATTTGCTTAAGAGAAATTAAGCAGACCCCAATATT 2520
QY 2521 GTAGTCAATTAATGATGATCAAAATTAAGAAATTTGTAACACAGAAATTAATT 2580
DB 2521 GTAGTCAATTAATGATGATCAAAATTAAGAAATTTGTAACACAGAAATTAATT 2580
QY 2581 AGAGAGATTCCTGCTGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 2618
DB 2581 AGAGAGATTCCTGCTGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 2618

RESULT 4
US-09-721-588-4370
Sequence 4370, Application US/09721588
GENERAL INFORMATION:
APPLICANT: Geating, David P.
APPLICANT: Holzman, Douglas A.
APPLICANT: Villaval, Jean-Luc
TITLE OF INVENTION: NOVEL NUCLEIC ACID MOLECULES AND USES
FILE REFERENCE: 1600.2046-001
CURRENT APPLICATION NUMBER: US/09/721.588
CURRENT FILING DATE: 2000-11-22
PRIOR APPLICATION NUMBER: 60/167,381
PRIOR FILING DATE: 1999-11-24
NUMBER OF SEQ ID NOS: 5410
SOFTWARE: FastSeq for Windows Version 4.0
SEQ ID NO 4370
LENGTH: 1453
TYPE: DNA
ORGANISM: Homo sapiens
US-09-721-588-4370

Query Match 49.6%; Score 1298; DB 32; Length 1453;
Best Local Similarity 100.0%; Pred. No. 6,78-271;
Matches 1298; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 466 GGGTCATGAGATCGGGGTTTCTGAGCATGATTCACCATCCCACTTTGGGTACAGAAC 525
DB 156 GGGTCATGAGATCGGGGTTTCTGAGCATGATTCACCATCCCACTTTGGGTACAGAAC 525
QY 526 TGACACCAATCAACGAG 585
DB 216 TGACACCAATCAACGAG 585
QY 586 GGGTCATGAGATCGGGGTTTCTGAGCATGATTCACCATCCCACTTTGGGTACAGAAC 645
DB 276 GGGTCATGAGATCGGGGTTTCTGAGCATGATTCACCATCCCACTTTGGGTACAGAAC 645
QY 646 TGGGCTGCGCATGCGGAG 705
DB 336 TGGGCTGCGCATGCGGAG 705
QY 706 ACTTCCTCTTCTTGAAGGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 765
DB 396 ACTTCCTCTTCTTGAAGGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 765
QY 766 ATCCCATCTCAAAATCTCAAGTCTGATGATGACCTTTCCCTAATTAAAGGCTTAAGCA 825
DB 456 ATCCCATCTCAAAATCTCAAGTCTGATGATGACCTTTCCCTAATTAAAGGCTTAAGCA 825
QY 826 TGGTGAAGGCGCATGAG 885
DB 516 TGGTGAAGGCGCATGAG 885
QY 886 GCGGCGGCGGCGAGATCACTGATCGATGATGATGATGATGATGATGATGATGATGAT 945
DB 576 GCGGCGGCGGCGAGATCACTGATCGATGATGATGATGATGATGATGATGATGATGAT 945


```

Qy 946 TCGGAGTATCTGAGAGTATGTTCTGACTTCTGTTAGTGTGCTGATTCGTGTT 1005
Db 636 TCGGAGTATCTGAGAGTATGTTCTGACTTCTGTTAGTGTGCTGATTCGTGTT 695
Qy 1006 GGTGTGAACGTCAGATTTCAATCAATCGCGTGGCTGTTTATGTTGTTCTCT 1065
Db 696 GGTGTGAACGTCAGATTTCAATCAATCGCGTGGCTGTTTATGTTGTTCTCT 755
Qy 1066 GTGTGTCAAGCTGTGCTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGT 1125
Db 756 GTGTGTCAAGCTGTGCTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGT 815
Qy 1126 CCAAGCTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGT 1185
Db 816 CCAAGCTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGT 875
Qy 1186 TTGGCATTCAGTGGGCTGTTTTCAGAGATCCAGCTGTGATTTGAAAGTCTTATTTGTC 1245
Db 876 TTGGCATTCAGTGGGCTGTTTTCAGAGATCCAGCTGTGATTTGAAAGTCTTATTTGTC 935
Qy 1246 ATGTGCAATCTAGTTTCTGTTTCTGTTTCTGTTTCTGTTTCTGTTTCTGTTTCT 1305
Db 936 ATGTGCAATCTAGTTTCTGTTTCTGTTTCTGTTTCTGTTTCTGTTTCTGTTTCT 995
Qy 1306 ACTTCTTGTGGGCTGTTTGTGGAGGCTCAAAATAGGCAAGACCTGAAAGCTGTTCTCC 1365
Db 996 ACTTCTTGTGGGCTGTTTGTGGAGGCTCAAAATAGGCAAGACCTGAAAGCTGTTCTCC 1055
Qy 1366 AGAGGGCTGTGAGAGACAGCCTGAGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGT 1425
Db 1056 AGAGGGCTGTGAGAGACAGCCTGAGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGT 1115
Qy 1426 CCTGTGAGCTGTGCGGAAGCAGATTGAGAGAGTGTGTGTGTGTGTGTGTGTGTGT 1485
Db 1116 CCTGTGAGCTGTGCGGAAGCAGATTGAGAGAGTGTGTGTGTGTGTGTGTGTGTGT 1175
Qy 1486 GGAATTTGAGAGCAATGCTGCTGCAACCTGTGCAATTTATGCAATTTTCTTAACT 1545
Db 1176 GGAATTTGAGAGCAATGCTGCTGCAACCTGTGCAATTTATGCAATTTTCTTAACT 1235
Qy 1546 TCTGCTCAGAAATGTCTCAATGTCTCTCAAGGTCTTCAATGTGTATCTTAACT 1605
Db 1236 TCTGCTCAGAAATGTCTCAATGTCTCTCAAGGTCTTCAATGTGTATCTTAACT 1295
Qy 1606 GACAGTTGCAAGTTTCAACCAATGAAAGCATTAGTCTGACAGTACAAATGTTGATCTC 1665
Db 1296 GACAGTTGCAAGTTTCAACCAATGAAAGCATTAGTCTGACAGTACAAATGTTGATCTC 1355
Qy 1666 CTGTGATTTACCAATATCTTCCCTGTATCTGTGCACTGAATCTTCTCACTGAACAT 1725
Db 1356 CTGTGATTTACCAATATCTTCCCTGTATCTGTGCACTGAATCTTCTCACTGAACAT 1415
Qy 1726 TTTTCTGCACTTTCAATGTATATAAAGAGTGTCTGT 1763
Db 1416 TTTTCTGCACTTTCAATGTATATAAAGAGTGTCTGT 1453

```

```

RESULT 5
US-09-634-754-740
; Sequence 740, Application US/09634754
; GENERAL INFORMATION:
; APPLICANT: Lal, Preeti
; APPLICANT: Au-Young, Janice
; APPLICANT: Guegler, Karl J.
; APPLICANT: Baughn, Mariah R.
; APPLICANT: Shah, Puryi
; TITLE OF INVENTION: DETECTING EXPRESSION OF GENES ENCODING SECRETED PROTEINS
; FILE REFERENCE: PA-0017 US
; CURRENT APPLICATION NUMBER: US/09/634,754
; NUMBER OF SEQ ID NOS: 809
; SOFTWARE: PERL Program

```

```

; SEQ ID NO 740
; LENGTH: 1557
; TYPE: DNA
; ORGANISM: Homo sapiens
; FEATURES:
; NAME/KEY: unsure
; LOCATION: 287, 299, 301, 305, 325, 329-330, 350-351, 361, 366, 371, 377
; NAME/KEY: unsure
; LOCATION: 395, 399, 407
; OTHER INFORMATION: a or g or c or t, unknown, or other
; NAME/KEY: misc feature
; OTHER INFORMATION: Incyte ID No: 1909132CT1
; PUBLICATION INFORMATION:
US-09-634-754-740

Query Match 43.8%; Score 1147; DB 27; Length 1557;
Best Local Similarity 100.0%; Pred. No. 2, 5e-238;
Matches 1147; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 766 ATCCATCTCCAAATCTGATGCTGCTGTGATGACCTTCCCTATCTTATAGCCCTAAGCA 825
Db 411 ATCCATCTCCAAATCTGATGCTGCTGTGATGACCTTCCCTATCTTATAGCCCTAAGCA 470
Qy 826 TGCTGAGCGCCATCAGACACGAGCGCTGCTGTCCATCTGTGSCCATCTGTACACT 885
Db 471 TGCTGAGCGCCATCAGACACGAGCGCTGCTGTCCATCTGTGSCCATCTGTACACT 530
Qy 886 GCCGCCGCCCAATACCTGTATCTGATGTGTGTGTGTGTGTGTGTGTGTGTGTGTGT 945
Db 531 GCCGCCGCCCAATACCTGTATCTGATGTGTGTGTGTGTGTGTGTGTGTGTGTGTGT 590
Qy 946 TCGGAGTATCTGAGAGTATGTTCTGACTTCTGTTAGTGTGCTGATTCGTGTT 1005
Db 591 TCGGAGTATCTGAGAGTATGTTCTGACTTCTGTTAGTGTGCTGATTCGTGTT 650
Qy 1006 GGTGTGAACGTCAGATTTCAATCAATCGCGTGGCTGTTTATGTTGTTCTCT 1065
Db 651 GGTGTGAACGTCAGATTTCAATCAATCGCGTGGCTGTTTATGTTGTTCTCT 710
Qy 1066 GTGTGTCAAGCTGTGCTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGT 1125
Db 711 GTGTGTCAAGCTGTGCTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGT 770
Qy 1126 CCAAGCTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGT 1185
Db 771 CCAAGCTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGT 830
Qy 1186 TTGCAATTCAGTGGGCTGTTTTCAGAGATCCAGCTGTGATTTGAAAGTCTTATTTGTC 1245
Db 831 TTGCAATTCAGTGGGCTGTTTTCAGAGATCCAGCTGTGATTTGAAAGTCTTATTTGTC 890
Qy 1246 ATGTGCAATCTAGTTTCTGTTTCTGTTTCTGTTTCTGTTTCTGTTTCTGTTTCT 1305
Db 891 ATGTGCAATCTAGTTTCTGTTTCTGTTTCTGTTTCTGTTTCTGTTTCTGTTTCT 950
Qy 1306 ACTTCTTGTGGGCTGTTTGTGGAGGCTCAAAATAGGCAAGACCTGAAAGCTGTTCTCC 1365
Db 951 ACTTCTTGTGGGCTGTTTGTGGAGGCTCAAAATAGGCAAGACCTGAAAGCTGTTCTCC 1010
Qy 1366 AGAGGGCTGTGAGAGACAGCCTGAGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGT 1425
Db 1011 AGAGGGCTGTGAGAGACAGCCTGAGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGT 1070
Qy 1426 CCTGTGAGCTGTGCGGAAGCAGATTGAGAGAGTGTGTGTGTGTGTGTGTGTGTGT 1485
Db 1071 CCTGTGAGCTGTGCGGAAGCAGATTGAGAGAGTGTGTGTGTGTGTGTGTGTGTGT 1130
Qy 1486 GGAATTTGAGAGCAATGCTGCCCTGCAACCTGGAATTAATGCAATTTTCTTAACT 1545
Db 1131 GGAATTTGAGAGCAATGCTGCCCTGCAACCTGGAATTAATGCAATTTTCTTAACT 1190
Qy 1546 TCTGCTCAGAAATGTCTCAATGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGT 1605

```

Db 1191 TCTGCTCCAGAAATGTCAGTGTGCTCCAGAGTCTTCCAAATAGATGTTATCTAACCT 1250
Qy 1606 GACAGTTGACAGTTTTCACCCATGAAAGCATTAAGTCTGACAGTACATGTTGATTC 1665
Db 1251 GACAGTTGACAGTTTTCACCCATGAAAGCATTAAGTCTGACAGTACATGTTGATTC 1310
Qy 1666 CTGATATTACCAATACATTTTCCCTGTTATCTTGACAGTAACTCTTCTACTGAACT 1725
Db 1311 CTGATATTACCAATACATTTTCCCTGTTATCTTGACAGTAACTCTTCTACTGAACT 1370
Qy 1726 TTTTCTGACATTTTCAATGTTATTAAGAGTGTGCTGTCACAACTTAACTCTTCT 1785
Db 1371 TTTTCTGACATTTTCAATGTTATTAAGAGTGTGCTGTCACAACTTAACTCTTCT 1430
Qy 1786 TATACCTGTTCCATCTGATAGTATCAAAAAGAAATTCCTTATTAATCTGTCAGACT 1845
Db 1431 TATACCTGTTCCATCTGATAGTATCAAAAAGAAATTCCTTATTAATCTGTCAGACT 1490
Qy 1846 ATGTTCCCTGAAATCAATGTTCCCTTTATGACCTGAGGCAATTAAGTGAAGCT 1905
Db 1491 ATGTTCCCTGAAATCAATGTTCCCTTTATGACCTGAGGCAATTAAGTGAAGCT 1550
Qy 1906 CAATTC 1912
Db 1551 CAATTC 1557

RESULT 6

US-09-634-754C-740

Sequence 740, Application US/09634754C
GENERAL INFORMATION:
APPLICANT: Lal, Preeti.
APPLICANT: Au-Young, Janice
APPLICANT: Guegler, Karl J.
APPLICANT: Baughn, Mariah R.
APPLICANT: Shah, Puri
TITLE OF INVENTION: DETECTING EXPRESSION OF GENES ENCODING SECRETED PROTEINS
FILE REFERENCE: PA-0017 US
CURRENT APPLICATION NUMBER: US/09/634,754C
CURRENT FILING DATE: 2002-04-26
NUMBER OF SEQ ID NOS: 809
SOFTWARE: PERL Program
SEQ ID NO 740
LENGTH: 1557
TYPE: DNA
ORGANISM: Homo sapiens
NAME/KEY: unsure
LOCATION: 287, 299, 301, 305, 325, 329-330, 350-351, 361, 366, 371, 377
FEATURE:
NAME/KEY: unsure
LOCATION: 395, 399, 407
OTHER INFORMATION: a or g or c or t, unknown, or other
FEATURE:
NAME/KEY: misc feature
OTHER INFORMATION: Inocyte ID No: 1909132CT1
PUBLICATION INFORMATION:
US-09-634-754C-740

Query Match 43.8%; Score 1147; DB 27; Length 1557;
Best Local Similarity 100.0%; Pred. No. 2.5e-238;
Matches 1147; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 766 ATCCATCTCCAAATCCCTGATGCTGATGACCTTTCCCTCACTTATAGGCTTAAGA 825
Db 411 ATCCATCTCCAAATCCCTGATGCTGATGACCTTTCCCTCACTTATAGGCTTAAGA 470
Qy 826 TGTGAGGCGCATGACGACCGAGCGCTGCTGCTCACTCTGAGCCCATCTGATCACT 885
Db 471 TGTGAGGCGCATGACGACCGAGCGCTGCTGCTCACTCTGAGCCCATCTGATCACT 530
Qy 886 GCGGCGCGCGCATGATCTGATGATGATGATGATGATGATGATGATGATGATGATGAT 945

Db 531 GCGGCGCGCGCATGATCTGATGATGATGATGATGATGATGATGATGATGATGATGAT 590
Qy 946 TGCAGATATCTGAGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 1005
Db 591 TGCAGATATCTGAGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 650
Qy 1006 GGTGGAACGTCAGATTTTCAATACATGCGGTGCTGCTTTTATAGTGTGCTCT 1065
Db 651 GGTGGAACGTCAGATTTTCAATACATGCGGTGCTGCTTTTATAGTGTGCTCT 710
Qy 1066 GTGGGTCACGCTGATGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 1125
Db 711 GTGGGTCACGCTGATGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 770
Qy 1126 CCAGGCTGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 1185
Db 771 CCAGGCTGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 830
Qy 1186 TTGGCATTCAGTGGGCTGCTTTTCCAGATTCACCTGATTTGGAAGTCTTATTTGTC 1245
Db 831 TTGGCATTCAGTGGGCTGCTTTTCCAGATTCACCTGATTTGGAAGTCTTATTTGTC 890
Qy 1246 ATGTCATCTAGTTTCAATTTTCTGTCGCTCTTAAAGAGTGCACACCATCATTT 1305
Db 891 ATGTCATCTAGTTTCAATTTTCTGTCGCTCTTAAAGAGTGCACACCATCATTT 950
Qy 1306 ACTTCTGCTGGGCTCTTTAGGACGCTCAAAATAGGACAACTGAAGTGTCTCC 1365
Db 951 ACTTCTGCTGGGCTCTTTAGGACGCTCAAAATAGGACAACTGAAGTGTCTCC 1010
Qy 1366 AAGAGGCTCTGAGACACGCTGATGATGATGATGATGATGATGATGATGATGATGAT 1425
Db 1011 AAGAGGCTCTGAGACACGCTGATGATGATGATGATGATGATGATGATGATGATGAT 1070
Qy 1426 CCTGAGCTGCTGGGAGAGAGATGATGATGATGATGATGATGATGATGATGATGATGAT 1485
Db 1071 CCTGAGCTGCTGGGAGAGAGATGATGATGATGATGATGATGATGATGATGATGATGAT 1130
Qy 1486 GAGACTTGAAGCAATGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 1545
Db 1131 GAGACTTGAAGCAATGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 1190
Qy 1546 TCTGCTCAGAAATGCTCAGTGTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 1605
Db 1191 TCTGCTCAGAAATGCTCAGTGTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 1250
Qy 1606 GACAGTTGATTTTCAAGCATGAAAGCATTAAGTCTGACAGTACATGTTGATTC 1665
Db 1251 GACAGTTGATTTTCAAGCATGAAAGCATTAAGTCTGACAGTACATGTTGATTC 1310
Qy 1666 CTGATATTACCAATACATTTTCCCTGTTATCTTGACAGTAACTCTTCTACTGAACT 1725
Db 1311 CTGATATTACCAATACATTTTCCCTGTTATCTTGACAGTAACTCTTCTACTGAACT 1370
Qy 1726 TTTTCTGACATTTTCAATGTTATTAAGAGTGTGCTGTCACAACTTAACTCTTCT 1785
Db 1371 TTTTCTGACATTTTCAATGTTATTAAGAGTGTGCTGTCACAACTTAACTCTTCT 1430
Qy 1786 TATACCTGTTCCATCTGATAGTATCAAAAAGAAATTCCTTATTAATCTGTCAGACT 1845
Db 1431 TATACCTGTTCCATCTGATAGTATCAAAAAGAAATTCCTTATTAATCTGTCAGACT 1490
Qy 1846 ATGTTCCCTGAAATCAATGTTCCCTTTATGACCTGAGGCAATTAAGTGAAGCT 1905
Db 1491 ATGTTCCCTGAAATCAATGTTCCCTTTATGACCTGAGGCAATTAAGTGAAGCT 1550
Qy 1906 CAATTC 1912
Db 1551 CAATTC 1557

RESULT 7

US-09-634-754D-740

```
; Sequence 740, Application US/09634754D
; GENERAL INFORMATION:
; APPLICANT: Lal, Preeti
; APPLICANT: Au-Young, Janice
; APPLICANT: Guegler, Karl J.
; APPLICANT: Baughn, Mariah R.
; APPLICANT: Shah, Puri
; TITLE OF INVENTION: DETECTING EXPRESSION OF GENES ENCODING SECRETED PROTEINS
; FILE REFERENCE: PA-0017 US
; CURRENT APPLICATION NUMBER: US/09/634,754D
; CURRENT FILING DATE: 2002-04-22
; NUMBER OF SEQ ID NOS: 809
; SOFTWARE: PERL Program
; SEQ ID NO 740
; LENGTH: 1557
; TYPE: DNA
; ORGANISM: Homo sapiens
; FEATURE:
; NAME/KEY: unsure
; LOCATION: 287, 299, 301, 305, 325, 329-330, 350-351, 361, 366, 371, 377
; FEATURE:
; NAME/KEY: unsure
; LOCATION: 395, 399, 407
; OTHER INFORMATION: a or g or c or t, unknown, or other
; FEATURE:
; NAME/KEY: misc_feature
; OTHER INFORMATION: Incyte ID No: 1909132CT1
; PUBLICATION INFORMATION:
; US-09-634-754D-740

Query Match      43.8%; Score 1147; DB 27; Length 1557;
Best Local Similarity 100.0%; Pred. No. 2,5e-238;
Matches 1147; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 766 ATCCCATCTCCAAATCTCTAGTCTGTGATGACCTTCTCTTATAGGCTTAAGCA 825
Db 411 ATCCCATCTCCAAATCTCTAGTCTGTGATGACCTTCTCTTATAGGCTTAAGCA 470

Qy 826 TGCAGACCGCATGACGACGAGCGGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 885
Db 471 TGCAGACCGCATGACGACGAGCGGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 530

Qy 886 GCCGCCGCCGCCGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 945
Db 531 GCCGCCGCCGCCGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 590

Qy 946 TGCAGATCTCTGAGTGTGATGATGATGATGATGATGATGATGATGATGATGATGAT 1005
Db 591 TGCAGATCTCTGAGTGTGATGATGATGATGATGATGATGATGATGATGATGATGAT 650

Qy 1006 GGTGTGAAAGTCAATTTCAATTCGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 1065
Db 651 GGTGTGAAAGTCAATTTCAATTCGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 710

Qy 1066 GTGGGTCAGGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 1125
Db 711 GTGGGTCAGGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 770

Qy 1126 CCAAGGCTGTAAGTCAATCTCTCAAGTGTGCTGCTGCTGCTGCTGCTGCTGCTGCT 1185
Db 771 CCAAGGCTGTAAGTCAATCTCTCAAGTGTGCTGCTGCTGCTGCTGCTGCTGCTGCT 830

Qy 1186 TTGGCATTCAGTGGGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 1245
Db 831 TTGGCATTCAGTGGGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 890

Qy 1246 ATGTCATCTAGTTCATTTCTGTCGCTCTTAACAGAGTGCACACCCCATCTTT 1305
Db 891 ATGTCATCTAGTTCATTTCTGTCGCTCTTAACAGAGTGCACACCCCATCTTT 950

Qy 1306 ACTTCTTGTGGGCTCTCTTTAGGAGGCTCAAAAATAGGACAGACCTGAAGCTGCTCC 1365
Db 951 ACTTCTTGTGGGCTCTCTTTAGGAGGCTCAAAAATAGGACAGACCTGAAGCTGCTCC 1010
```

```
Qy 1366 AGAGGCTCTGACAGACAGCCTGAGGTGATGAAGGTGAGGCTGCTCTCAGAAA 1425
Db 1011 AGAGGCTCTGACAGACAGCCTGAGGTGATGAAGGTGAGGCTGCTCTCAGAAA 1070

Qy 1426 CCCTGAGCTGTGCGGAAGCAGATTGAGCAGTGAAGGAAGAACTTGCCTCTCAGCA 1485
Db 1071 CCCTGAGCTGTGCGGAAGCAGATTGAGCAGTGAAGGAAGAACTTGCCTCTCAGCA 1130

Qy 1486 GGATTTGAGAGCAATGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 1545
Db 1131 GGATTTGAGAGCAATGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 1190

Qy 1546 TCTGCTCAGAAATGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 1605
Db 1191 TCTGCTCAGAAATGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 1250

Qy 1606 GACAGTTGAGTTTCAACCCATGGAAGCAATTAAGTCAAGTCAATGTTGATTTCTC 1665
Db 1251 GACAGTTGAGTTTCAACCCATGGAAGCAATTAAGTCAAGTCAATGTTGATTTCTC 1310

Qy 1666 CTGATTTTACCAATATCTTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 1725
Db 1311 CTGATTTTACCAATATCTTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 1370

Qy 1726 TTTTCTGACATTTTCAATGATTAAGAGGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 1785
Db 1371 TTTTCTGACATTTTCAATGATTAAGAGGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 1430

Qy 1786 TATACCTGTTTCTTACCTGATGATGATGATGATGATGATGATGATGATGATGATGAT 1845
Db 1431 TATACCTGTTTCTTACCTGATGATGATGATGATGATGATGATGATGATGATGATGAT 1490

Qy 1846 ATGTTCCCTGAAATATGTTCCCTTTTATGATGAGGATTAATGCAATGCAATGCA 1905
Db 1491 ATGTTCCCTGAAATATGTTCCCTTTTATGATGAGGATTAATGCAATGCAATGCA 1550

Qy 1906 CAATTCT 1912
Db 1551 CAATTCT 1557

RESULT 8
US-10-745-444-740
; Sequence 740, Application US/10745444
; GENERAL INFORMATION:
; APPLICANT: Lal, Preeti
; APPLICANT: Au-Young, Janice
; APPLICANT: Guegler, Karl J.
; APPLICANT: Baughn, Mariah R.
; APPLICANT: Shah, Puri
; TITLE OF INVENTION: DETECTING EXPRESSION OF GENES ENCODING SECRETED PROTEINS
; FILE REFERENCE: PA-0017 US
; CURRENT APPLICATION NUMBER: US/10/745,444
; CURRENT FILING DATE: 2003-12-23
; NUMBER OF SEQ ID NOS: 809
; SOFTWARE: PERL Program
; SEQ ID NO 740
; LENGTH: 1557
; TYPE: DNA
; ORGANISM: Homo sapiens
; FEATURE:
; NAME/KEY: unsure
; LOCATION: 287, 299, 301, 305, 325, 329-330, 350-351, 361, 366, 371, 377
; FEATURE:
; NAME/KEY: unsure
; LOCATION: 395, 399, 407
; OTHER INFORMATION: a or g or c or t, unknown, or other
; FEATURE:
; NAME/KEY: misc_feature
; OTHER INFORMATION: Incyte ID No: 1909132CT1
; PUBLICATION INFORMATION:
; US-10-745-444-740
```


Db 241 GTTGTGTGTAAGTGTGATTTCAATTCAGTGTGTTTATATGTGTGTT 300
Qy 1062 CTCTGTGTGTGTAAGTGTGATTTCAATTCAGTGTGTTTATATGTGTGTT 1121
Db 301 CTCTGTGTGTGTAAGTGTGATTTCAATTCAGTGTGTTTATATGTGTGTT 360
Qy 1122 CTGACGAGGCTGTGATGTGATTCCTGTGATGTGATTCCTGTGATGTGAT 1181
Db 361 CTGACGAGGCTGTGATGTGATTCCTGTGATGTGATTCCTGTGATGTGAT 420
Qy 1182 CCTTTTGTGATGTGATGTGATTCCTGTGATGTGATTCCTGTGATGTGAT 1241
Db 421 CCTTTTGTGATGTGATGTGATTCCTGTGATGTGATTCCTGTGATGTGAT 480
Qy 1242 TGTGATGTGATGTGATGTGATTCCTGTGATGTGATTCCTGTGATGTGAT 1301
Db 481 TGTGATGTGATGTGATGTGATTCCTGTGATGTGATTCCTGTGATGTGAT 540
Qy 1302 ATTTACTTCTGTGATGTGATTCCTGTGATGTGATTCCTGTGATGTGAT 1361
Db 541 ATTTACTTCTGTGATGTGATTCCTGTGATGTGATTCCTGTGATGTGAT 600
Qy 1362 CTGACGAGGCTGTGATGTGATTCCTGTGATGTGATTCCTGTGATGTGAT 1421
Db 601 CTGACGAGGCTGTGATGTGATTCCTGTGATGTGATTCCTGTGATGTGAT 660
Qy 1422 GAAACCTGTGATGTGATGTGATTCCTGTGATGTGATTCCTGTGATGTGAT 1481
Db 661 GAAACCTGTGATGTGATGTGATTCCTGTGATGTGATTCCTGTGATGTGAT 720
Qy 1482 GACGAGCTGTGATGTGATTCCTGTGATGTGATTCCTGTGATGTGAT 1541
Db 721 GACGAGCTGTGATGTGATTCCTGTGATGTGATTCCTGTGATGTGAT 780
Qy 1542 GCTTCTGTGATGTGATGTGATTCCTGTGATGTGATTCCTGTGATGTGAT 1601
Db 781 GCTTCTGTGATGTGATGTGATTCCTGTGATGTGATTCCTGTGATGTGAT 840
Qy 1602 ACCGTGAGTGTGATGTGATTCCTGTGATGTGATTCCTGTGATGTGAT 1661
Db 841 ACCGTGAGTGTGATGTGATTCCTGTGATGTGATTCCTGTGATGTGAT 900
Qy 1662 TCTCTGTGATGTGATGTGATTCCTGTGATGTGATTCCTGTGATGTGAT 1721
Db 901 TCTCTGTGATGTGATGTGATTCCTGTGATGTGATTCCTGTGATGTGAT 960
Qy 1722 CACTTCTGTGATGTGATGTGATTCCTGTGATGTGATTCCTGTGATGTGAT 1781
Db 961 CACTTCTGTGATGTGATGTGATTCCTGTGATGTGATTCCTGTGATGTGAT 1020
Qy 1782 TCTTATACCTGTGATGTGATTCCTGTGATGTGATTCCTGTGATGTGAT 1841
Db 1021 TCTTATACCTGTGATGTGATTCCTGTGATGTGATTCCTGTGATGTGAT 1080
Qy 1842 GACTATGTGATGTGATGTGATTCCTGTGATGTGATTCCTGTGATGTGAT 1901
Db 1081 GACTATGTGATGTGATGTGATTCCTGTGATGTGATTCCTGTGATGTGAT 1140
Qy 1902 AGCTCA 1907
Db 1141 AGCTCA 1146

RESULT 10
US-09-867-570-3
; Sequence 3, Application US/09867570
; GENERAL INFORMATION:
; APPLICANT: MEI, Ming-Hui et al.
; TITLE OF INVENTION: ISOLATED HUMAN G-PROTEIN COUPLED
; TITLE OF INVENTION: RECEPTORS, NUCLEIC ACID MOLECULES ENCODING HUMAN GPCR
; FILE REFERENCE: CLO00900-CIP
; CURRENT APPLICATION NUMBER: US/09/867,570

; CURRENT FILING DATE: 2001-05-31
; PRIOR APPLICATION NUMBER: 09/695,045
; PRIOR FILING DATE: 2000-10-25
; NUMBER OF SEQ ID NOS: 4
; SOFTWARE: FASTSEQ for Windows Version 4.0
; SEQ ID NO 3
; LENGTH: 8622
; TYPE: DNA
; ORGANISM: Human
US-09-867-570-3

Query Match 42.1%; Score 1101; DB 36; Length 8622;
Best Local Similarity 100.0%; Pred. No. 1,3e-228;
Matches 1101; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 466 GGGTCATCAGACGAGGCTTTCTGAGCATGTGATTCACATCCCACTGGGTACAGAAC 525
Db 7502 GGGTCATCAGACGAGGCTTTCTGAGCATGTGATTCACATCCCACTGGGTACAGAAC 7561
Qy 526 TGACCAATCAACGAGAGGAGAGACTCTTGTCTCAAGCAGACCTGAGCTTCAAG 585
Db 7562 TGACCAATCAACGAGAGGAGAGACTCTTGTCTCAAGCAGACCTGAGCTTCAAG 7621
Qy 586 GGGTCATCAGACGAGGCTTTCTGAGCATGTGATTCACATCCCACTGGGTACAGAAC 645
Db 7622 GGGTCATCAGACGAGGCTTTCTGAGCATGTGATTCACATCCCACTGGGTACAGAAC 7681
Qy 646 TGGGCTCCGCTGAGAGGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 705
Db 7682 TGGGCTCCGCTGAGAGGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 7741
Qy 706 ACTTCTCTCTCTTCTGAGGAGGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 765
Db 7742 ACTTCTCTCTCTTCTGAGGAGGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 7801
Qy 766 ATCCCATCTCCAAATCTCTGAGTGTGATTCACATCCCACTGGGTACAGAAC 825
Db 7802 ATCCCATCTCCAAATCTCTGAGTGTGATTCACATCCCACTGGGTACAGAAC 7861
Qy 826 TGTGAGAGGCTATGAG 885
Db 7862 TGTGAGAGGCTATGAG 7921
Qy 886 GCGCGCGCCGAGATCTCTGATGTGATTCACATCCCACTGGGTACAGAAC 945
Db 7922 GCGCGCGCCGAGATCTCTGATGTGATTCACATCCCACTGGGTACAGAAC 7981
Qy 946 TGTGAGATCTCTGAGTGTGATTCACATCCCACTGGGTACAGAAC 1005
Db 7982 TGTGAGATCTCTGAGTGTGATTCACATCCCACTGGGTACAGAAC 8041
Qy 1006 GGTGTGAGAGGCTATGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 1065
Db 8042 GGTGTGAGAGGCTATGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 8101
Qy 1066 GTGGGTCAAGCTGTGAGTGTGATTCACATCCCACTGGGTACAGAAC 1125
Db 8102 GTGGGTCAAGCTGTGAGTGTGATTCACATCCCACTGGGTACAGAAC 8161
Qy 1126 CCAAGCTGTGAGTGTGATTCACATCCCACTGGGTACAGAAC 1185
Db 8162 CCAAGCTGTGAGTGTGATTCACATCCCACTGGGTACAGAAC 8221
Qy 1186 TTTGATTCAGTGTGAGTGTGATTCACATCCCACTGGGTACAGAAC 1245
Db 8222 TTTGATTCAGTGTGAGTGTGATTCACATCCCACTGGGTACAGAAC 8281
Qy 1246 ATGTGATTCAGTGTGATTCACATCCCACTGGGTACAGAAC 1305
Db 8282 ATGTGATTCAGTGTGATTCACATCCCACTGGGTACAGAAC 8341
Qy 1306 ACTTCTCTCTCTTCTGAGGAGGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 1365

Db 8342 ACTTCTTGTGGGCTCTTTAGGACGCTGCAAAATAGGAGAACTGAAAGCTGTTCTCC 8401
Qy 1366 AGAGGGCTCTGAGGACACGCTGAGGTGATGAGGTGAGGTGGCTTCTCAGAGAA 1425
Db 8402 AGAGGGCTCTGAGGACACGCTGAGGTGATGAGGTGAGGTGGCTTCTCAGAGAA 8461
Qy 1426 CCTGAGAGCTGTGGGAAAGCAGATTGAGCAAGTGAAGAAACCTTCTGCTGTGACAGA 1485
Db 8462 CCTGAGAGCTGTGGGAAAGCAGATTGAGCAAGTGAAGAAACCTTCTGCTGTGACAGA 8521
Qy 1486 GGACTTTAGAGCAATGCTGCTGCTGCAACCTTGAACATTAATATGATTTTCTTAGCCT 1545
Db 8522 GGACTTTAGAGCAATGCTGCTGCTGCAACCTTGAACATTAATATGATTTTCTTAGCCT 8581
Qy 1546 TCTGCTCAGAAATGTCTCAG 1566
Db 8582 TCTGCTCAGAAATGTCTCAG 8602

RESULT 11
US-60-201-715-47/c
; Sequence 47, Application US/60201715
; GENERAL INFORMATION:
; APPLICANT: Beasley, Ellen
; TITLE OF INVENTION: ISOLATED HUMAN G-PROTEIN COUPLED
; TITLE OF INVENTION: RECEPTORS, NUCLEIC ACID MOLECULES ENCODING HUMAN GPCR
; FILE REFERENCE: C100506
; CURRENT APPLICATION NUMBER: US/60/201,715
; CURRENT FILING DATE: 2000-05-03
; NUMBER OF SEQ ID NOS: 124
; SOFTWARE: PasteSeq for Windows Version 4.0
; SEQ ID NO 47
; LENGTH: 3988
; TYPE: DNA
; ORGANISM: HUMAN
US-60-201-715-47

Query Match 41.9%; Score 1098; DB 85; Length 3988;
Best Local Similarity 100.0%; Pred. No. 7,1e-228;
Matches 1098; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 466 GGGTCATCAGACTGGGGTTTCTGAGCATTGATTCACATCCAGTCTTGGGTACAGAAC 525
Db 1173 GGGTCATCAGACTGGGGTTTCTGAGCATTGATTCACATCCAGTCTTGGGTACAGAAC 1114
Qy 526 TGAACCAATGACAGAGCTGAGAGACCTTCTCAACAGAGACCCCTGAGCTTCAAGG 585
Db 1113 TGAACCAATGACAGAGCTGAGAGACCTTCTCAACAGAGACCCCTGAGCTTCAAGG 1054
Qy 586 GGCTGACGTGATCGTTTCCCTTGTGCGCTGACAGAGAAACGGGTGTGCTCTGCTCC 645
Db 1053 GGCTGACGTGATCGTTTCCCTTGTGCGCTGACAGAGAAACGGGTGTGCTCTGCTCC 994
Qy 646 TGGGCTGCGCAGTGGCAGAGAAAGCTGTCTCAATCAATCTCAACCTGCTGCGCGCG 705
Db 993 TGGGCTGCGCAGTGGCAGAGAAAGCTGTCTCAATCAATCTCAACCTGCTGCGCGCG 934
Qy 706 ACTTCTCTTCTTCTTGGGCGGCAATTATATGTTGCGCGTTAGCGCCATCAATATCGCGC 765
Db 933 ACTTCTCTTCTTCTTGGGCGGCAATTATATGTTGCGCGTTAGCGCCATCAATATCGCGC 874
Qy 766 ATCCATCTCCAAAATCTCAGTCTGTGATGACCTTTCCTCACTTATATAGGCTTAAGA 825
Db 873 ATCCATCTCCAAAATCTCAGTCTGTGATGACCTTTCCTCACTTATATAGGCTTAAGA 814
Qy 826 TGTGAGGCGCATCAGACCGAGCGCTGCTGTCAATCTGTGAGCCATCTGGTACACT 885
Db 813 TGTGAGGCGCATCAGACCGAGCGCTGCTGTCAATCTGTGAGCCATCTGGTACACT 754
Qy 886 GCGCGCGCGCGCATCAGTCTGTGATGATGTTGCTCTGCTGAGGCGCTGCTGCTGCTG 945
Db 753 GCGCGCGCGCGCATCAGTCTGTGATGATGTTGCTCTGCTGAGGCGCTGCTGCTGCTG 694

Qy 946 TCGGAGTATCTCTGAGTGAATGTTCTGTGACTTCTGTTTAGTGTGCTGATTTCTGTTT 1005
Db 693 TCGGAGTATCTCTGAGTGAATGTTCTGTGACTTCTGTTTAGTGTGCTGATTTCTGTTT 634
Qy 1006 GGTGTGAAACGTCAATTTTCAATTAATTCGCTGCTGCTGCTGCTGCTGCTGCTGCT 1065
Db 633 GGTGTGAAACGTCAATTTTCAATTAATTCGCTGCTGCTGCTGCTGCTGCTGCTGCT 574
Qy 1066 GTGGGTCCAGCTGGTCTGCTGTGTGAGATTCCTGTGTGATTCGCGAAGATGCGCTGA 1125
Db 573 GTGGGTCCAGCTGGTCTGCTGTGTGAGATTCCTGTGTGATTCGCGAAGATGCGCTGA 514
Qy 1126 CCAGGCTGTAGCTGACCAATCTCTCTCAGAGTGTGCTTCTCTGTGAGCTGCTCCCT 1185
Db 513 CCAGGCTGTAGCTGACCAATCTCTCTCAGAGTGTGCTTCTCTGTGAGCTGCTCCCT 454
Qy 1186 TTGGCATTAAGTGGGCTGCTGTTTTCAGAGATTCACACTGATTTGAAATCTTATTTTC 1245
Db 453 TTGGCATTAAGTGGGCTGCTGTTTTCAGAGATTCACACTGATTTGAAATCTTATTTTC 394
Qy 1246 ATGTGATCTAGTTTCAATTTTCTGTGCGCTCTTAACAGAGTSCCAACCCATCATTT 1305
Db 393 ATGTGATCTAGTTTCAATTTTCTGTGCGCTCTTAACAGAGTSCCAACCCATCATTT 334
Qy 1306 ACTTCTTGTGGGCTCTTTAGGAGCGTCAAAATAGGCAAACTGAAAGCTGTTCTCC 1365
Db 333 ACTTCTTGTGGGCTCTTTAGGAGCGTCAAAATAGGCAAACTGAAAGCTGTTCTCC 274
Qy 1366 AGAGGGCTCTGAGAGACAGCTGAGGTGATGAGAGGTGAGGTGCTTCTCAAGAGAA 1425
Db 273 AGAGGGCTCTGAGAGACAGCTGAGGTGATGAGAGGTGAGGTGCTTCTCAAGAGAA 214
Qy 1426 CCTGAGAGCTGTGCGGAAAGCAGATTGAGAGAGTGAAGAACTCTGCTGCTGACAGA 1485
Db 213 CCTGAGAGCTGTGCGGAAAGCAGATTGAGAGAGTGAAGAACTCTGCTGCTGACAGA 154
Qy 1486 GGAATTTGAGCAATGCTGCTGCTGCAACCTTGAACATTAATATGATTTTCTTAGCCT 1545
Db 153 GGAATTTGAGCAATGCTGCTGCTGCAACCTTGAACATTAATATGATTTTCTTAGCCT 94
Qy 1546 TCTGCTCAGAAATGTCT 1563
Db 93 TCTGCTCAGAAATGTCT 76

RESULT 12
US-10-292-798-1273
; Sequence 1273, Application US/10292798
; GENERAL INFORMATION:
; APPLICANT: SUMA, MAKIKO
; APPLICANT: ASAI, KIYOSHI
; APPLICANT: AKIYAMA, YUTAKA
; APPLICANT: ABURATANI, HIROYUKI
; TITLE OF INVENTION: GUANOSINE TRIPHOSPHATE-BINDING PROTEIN COUPLED RECEPTORS
; FILE REFERENCE: 08435/166
; CURRENT APPLICATION NUMBER: US/10/292,798
; PRIOR FILING DATE: 2002-11-13
; PRIOR APPLICATION NUMBER: 10/017,161
; PRIOR FILING DATE: 2001-12-18
; PRIOR APPLICATION NUMBER: JP 2001-246789
; NUMBER OF SEQ ID NOS: 2070
; SOFTWARE: Patentin Ver. 2.1
; SEQ ID NO 1273
; LENGTH: 1369
; TYPE: DNA
; ORGANISM: Homo sapiens
; FEATURE:
; LOCATION: source
; FEATURE:
; LOCATION: (1)..(1369)
; FEATURE:

```

; NAME/KEY: CDS
; LOCATION: (201) .. (1169)
US-10-292-798-1273

```

Query Match	41.7%	Score 1093;	DB 50;	Length 1369;
Best Local Similarity	99.8%;	Pred. No. 1.2e-226;		
Matches 1193; Conservative	0;	Mismatches 2;	Indels 0;	Gaps 0;

Qy	466	GGGTCATCAGCTGGGGGTTTCTTGAGCAGTGAATTCAAACATCCCACTCTTGAGTACAGAAC	525
Dp	175	GGGTCATCAGCTGGGGGTTTCTTGAGCAGTGAATTCAAACATCCCACTCTTGAGTACAGAAC	234
Qy	526	TGACACCAATCAACGAGCGTGAAGAGACTCTTGTCTCAAGACAGACCCCTGAGCTTACGCG	585
Dp	235	TGACACCAATCAACGAGCGTGAAGAGACTCTTGTCTCAAGACAGACCCCTGAGCTTACGCG	294
Qy	586	GGCTGACGTGCATCGTTTCCCTTGTGCGGCTGACAGAGAAAGCGGGTGTGCTGTGGCTCC	645
Dp	295	GGCTGACGTGCATCGTTTCCCTTGTGCGGCTGACAGAGAAAGCGGGTGTGCTGTGGCTCC	354
Qy	646	TGGGCTGCGCGATGGGAGAGACGGTGTCTCCATCTAATACCTCAACCTGGTCGCGGCG	705
Dp	355	TGGGCTGCGCGATGGGAGAGACGGTGTCTCCATCTAATACCTCAACCTGGTCGCGGCG	414
Qy	706	ACTTCTCTTTCCTTAGCGGCCACATTAATATGTTGCGCGCTAAGCCTCAATATCCGCG	765
Dp	415	ACTTCTCTTTCCTTAGCGGCCACATTAATATGTTGCGCGCTAAGCCTCAATATCCGCG	474
Qy	766	ATCCCATCTCCAAAATCCTCAGTCTGTGATGACCTTTCCCTACTTTATAGGCTTAAAGCA	825
Dp	475	ATCCCATCTCCAAAATCCTCAGTCTGTGATGACCTTTCCCTACTTTATAGGCTTAAAGCA	534
Qy	826	TGCTGAGGGCCATCAGACACGAGCGCTGCTGTCATCTGTGAGGCCATCTGTGACACT	885
Dp	535	TGCTGAGGGCCATCAGACACGAGCGCTGCTGTCATCTGTGAGGCCATCTGTGACACT	594
Qy	886	GCGCGCGGCCAGATACCTGTCAATCGTCAATGTGTGCTGTGCGGCGCTGTGCTGCG	945
Dp	595	GCGCGCGGCCAGATACCTGTCAATCGTCAATGTGTGCTGTGCGGCGCTGTGCTGCG	654
Qy	946	TGCGGAGATCCTGAGAGTGAATGTTCTGTGACTTCTCTGTAAAGTGTGTGATTTCTGTTT	1005
Dp	655	TGCGGAGATCCTGAGAGTGAATGTTCTGTGACTTCTCTGTGTAAAGTGTGTGATTTCTGTTT	714
Qy	1006	GGTGTGAAACGTCAGATTTCAATTACAATCGCGTGGCTGGTTTTTTATATGTGATTTCTCT	1065
Dp	715	GGTGTGAAACGTCAGATTTCAATTACAATCGCGTGGCTGGTTTTTTATATGTGATTTCTCT	774
Qy	1066	GTGGGTCCAGGCTGTGCTGTGCTGTGAGATTTCTGTGTGATCCCGAAGATGCGGCTGA	1125
Dp	775	GTGGGTCCAGGCTGTGCTGTGCTGTGAGATTTCTGTGTGATCCCGAAGATGCGGCTGA	834
Qy	1126	CCAGGCTGTAGTGAACCATCTCTCTCAACAGTGTGTGTCTTCTCTCTCTGTGGCTTGCCCT	1185
Dp	835	CCAGGCTGTAGTGAACCATCTCTCTCAACAGTGTGTGTCTTCTCTCTCTGTGGCTTGCCCT	894
Qy	1186	TTGGCATTCAGTGGGCGCTGTGTTTTTCAGGATTCACCTGGATTGGAAGATCTTATTTTGTG	1245
Dp	895	TTGGCATTCAGTGGGCGCTGTGTTTTTCAGGATTCACCTGGATTGGAAGATCTTATTTTGTG	954
Qy	1246	ATGTGCATCTAAGTTTCAATTTTCTGTGCGCTCTTAAACAGCAGTGCACCAACCCCATCATTT	1305
Dp	955	ATGTGCATCTAAGTTTCAATTTTCTGTGCGCTCTTAAACAGCAGTGCACCAACCCCATCATTT	101
Qy	1306	ACTTCTTGTGAGGCTCTTTTAGGCAAGGCTCAAAATAGCAGAAACCTGAAAGCTGTGTTCC	1366
Dp	1015	ACTTCTTGTGAGGCTCTTTTAGGCAAGGCTCAAAATAGCAGAAACCTGAAAGCTGTGTTCC	107
Qy	1366	AGAGGGCTCTCAGAGACACGCTGAGTGTGATGTAAGTGTGAGGGTGTGCTTCTCAGAGAA	1425
Dp	1075	AGAGGGCTCTCAGAGACACGCTGAGTGTGATGTAAGTGTGAGGGTGTGCTTCTCAGAGAA	113
Qy	1426	CCCTGAGCTGTGCGGAAGCAGATTTGAGACGTGAGAAAGAACTTCTGCGCTGTGACACA	1485

Accession	Sequence	Position
D8	CCCTGGAGCTGTGGGAAGCAGATTGGAGAGTAAAGAAAGCTCTGCGCTGCAGCA	1194
QY	GGACTTTGAGGCAATGCTGCGCTGCACACCTTGACATTATATGCAATTTTCTTAAGCT	1545
D8	GGAATTTGAGAGCAATGCTGCGCTGCACACCTTGACATTATATGCAATTTTCTTAAGCT	1254
QY	TCGCGCTCAGAAATGTCTCAGTGTGCTCCCAAGGCTCTTGAAATGATGTTATCTTAACCT	1605
D8	TCGCGCTCAGAAATGTCTCAGTGTGCTCCCAAGGCTCTTGAAATGATGTTATCTTAACCT	1314
QY	GACAGTTGCAATTTTCAACCATGGAAGCAATTACTGCAAGTACCAATGTTTGGAA	1660
D8	GACAGTTGCAATTTTCAACCATGGAAGCAATTACTGCAAGTACCAATGTTTGGAA	1369

```

RESULT 13
PCT-US01-14519-30
Sequence 30, Application PC/RUS0114519
GENERAL INFORMATION:
APPLICANT: California Institute of Technology
INVENTOR: Anderson, David J.
APPLICANT: Dong, Xinzhong
APPLICANT: Zylka, Mark
APPLICANT: Simon, Melvin
APPLICANT: Han, Sang-kyou
TITLE OF INVENTION: PAIR SIGNALING MOLECULES
FILE REFERENCE: CALTEC 004VPC
CURRENT APPLICATION NUMBER: PCT/US01/14519
CURRENT FILING DATE: 2001-05-04
PRIOR APPLICATION NUMBER: US 60/222,344
PRIOR FILING DATE: 2000-08-01
PRIOR APPLICATION NUMBER: US 60/202,027
PRIOR FILING DATE: 2000-05-04
PRIOR APPLICATION NUMBER: US 09/704,707
PRIOR FILING DATE: 2000-11-03
PRIOR APPLICATION NUMBER: US unknown
PRIOR FILING DATE: 2001-04-19
NUMBER OF SEQ ID NOS: 109
SOFTWARE: FastSeq for Windows Version 4.0
SEQ ID NO 30
LENGTH: 1400
TYPE: DNA
ORGANISM: Homo sapiens
FEATURES:
NAME/KEY: CDS
LOCATION: (332)...(1297)
PCT-US01-14519-30

```

Query Match	41.6%	Score 1088;	DB 1;	Length 1400;
Best Local Similarity	100.0%	Pred. No. 1.4e-225;		
Matches 1088; Conservative	0;	Mismatches	0;	Gaps 0;

Oy	473	CAGACTGGGGTTTCTGAGCATGAGATTCAACATCCAGTCTTGGGTACAGAACTGACACC	532
Db	313	CAGACTGGGGTTTCTGAGCATGAGATTCAACATCCAGTCTTGGGTACAGAACTGACACC	372
Oy	533	AATCAACGAGCGTAGAGAGACTCCTTGCTACAGACAGCCTTGACTTCAACGGGGCTGAC	592
Db	373	AATCAACGAGCGTAGAGAGACTCCTTGCTACAGACAGCCTTGACTTCAACGGGGCTGAC	432
Oy	593	GTGCATGTTTTCCCTTGCGCGCTGACAGAGAAAAGCGGGTTGTGCTTGAGGCTCCTGGGGCTG	652
Db	433	GTGCATGTTTTCCCTTGCGCGCTGACAGAGAAAAGCGGGTTGTGCTTGAGGCTCCTGGGGCTG	492
Oy	653	CCGCATGCGAGGAACGCTGTCTTCATCTACATCTTCAACCTGTGTGCGCGGCGCACTTCT	712
Db	493	CCGCATGCGAGGAACGCTGTCTTCATCTACATCTTCAACCTGTGTGCGCGGCGCACTTCT	552
Oy	713	CTTCTCTTAGCGGCGCAATTATATGTTTGGCGTTAAGCGCTATCAATATTCGCGCATTCCTAT	772
Db	553	CTTCTCTTAGCGGCGCAATTATATGTTTGGCGTTAAGCGCTATCAATATTCGCGCATTCCTAT	612


```

QY 773 CTCGAAATCTCTAGTCTGTGATGACCTTTCTTAAAGGCTTAAGACATGCTGAG 832
DB 613 CTCGAAATCTCTAGTCTGTGATGACCTTTCTTAAAGGCTTAAGACATGCTGAG 672
QY 833 CGCCATGACACCGGAGGCTGTGATGACCTTTCTTAAAGGCTTAAGACATGCTGAG 892
DB 673 CGCCATGACACCGGAGGCTGTGATGACCTTTCTTAAAGGCTTAAGACATGCTGAG 732
QY 893 CCCGATGACACCGGAGGCTGTGATGACCTTTCTTAAAGGCTTAAGACATGCTGAG 952
DB 733 CCCGATGACACCGGAGGCTGTGATGACCTTTCTTAAAGGCTTAAGACATGCTGAG 792
QY 953 TATCTGAGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 1012
DB 793 TATCTGAGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 852
QY 1013 AACGTCAGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 1072
DB 853 AACGTCAGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 912
QY 1073 CAGCTGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 1132
DB 913 CAGCTGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 972
QY 1133 GTACGTGACATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 1192
DB 973 GTACGTGACATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 1032
QY 1193 TCAATGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 1252
DB 1033 TCAATGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 1092
QY 1253 TCTAGTTTCAATTTCTCTGCTGCTCTTAAAGAGAGGAGGAGGAGGAGGAGGAGGAG 1312
DB 1093 TCTAGTTTCAATTTCTCTGCTGCTCTTAAAGAGAGGAGGAGGAGGAGGAGGAGGAG 1152
QY 1313 CGTGGGCTCTTTAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 1372
DB 1153 CGTGGGCTCTTTAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 1212
QY 1373 TCTGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 1432
DB 1213 TCTGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 1272
QY 1433 GCTGTGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 1492
DB 1273 GCTGTGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 1332
QY 1493 GAGAGCAATGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 1552
DB 1333 GAGAGCAATGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 1392
QY 1553 CAGAAATG 1560
DB 1393 CAGAAATG 1400

```

RESULT 14
PCT-US03-15004-30

```

; Sequence 30, Application PC/TUS0315004
; GENERAL INFORMATION:
; APPLICANT: California Institute of Technology
; APPLICANT: Anderson, David J.
; APPLICANT: Dong, Xinzhang
; APPLICANT: Zylka, Mark
; APPLICANT: Simon, Melvin
; APPLICANT: Han, Sang-kyou
; TITLE OF INVENTION: PAIN SIGNALING MOLECULES
; FILE REFERENCE: CALTE.4C10PC
; CURRENT APPLICATION NUMBER: PCT/US03/15004
; CURRENT FILING DATE: 2003-05-08
; PRIOR APPLICATION NUMBER: US 10/183,116
; PRIOR FILING DATE: 2002-06-26

```

```

; NUMBER OF SEQ ID NOS: 109
; SOFTWARE: FastSeq For Windows Version 4.0
; SEQ ID NO 30
; LENGTH: 1400
; TYPE: DNA
; ORGANISM: Homo sapiens
; NAME/KEY: CDS
; LOCATION: (332) ... (1297)
; PCT-US03-15004-30

```

```

Query Match 41.6%; Score 1088; DB 2; Length 1400;
Best Local Similarity 100.0%; Pred. No. 1.4e-225; Indels 0; Gaps 0;
Matches 1088; Conservative 0; Mismatches 0;

```

```

QY 473 CAGCTGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 532
DB 313 CAGCTGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 372
QY 533 AATCAACGAGCGTAGAGAGACTCTTGTCAAGAGACCTGAGCTTACGCGGCTGAC 592
DB 373 AATCAACGAGCGTAGAGAGACTCTTGTCAAGAGACCTGAGCTTACGCGGCTGAC 432
QY 593 GTGCAATGCTTCCCTTGTGCGGCTGAGCAGAGAAAGCGGCTGTGCTGCGGCTG 652
DB 433 GTGCAATGCTTCCCTTGTGCGGCTGAGCAGAGAAAGCGGCTGTGCTGCGGCTG 492
QY 653 CGGCATGAGCAGAGAGCGCTGCTCCATCTAATCTCAACCTGATCGCGGCGACTTCT 712
DB 493 CGGCATGAGCAGAGAGCGCTGCTCCATCTAATCTCAACCTGATCGCGGCGACTTCT 552
QY 713 CTTCCTTAAAGGAGCAGATTAATGTTGCGCGTTAGCGCTCAATATCCGCATCCAT 772
DB 553 CTTCCTTAAAGGAGCAGATTAATGTTGCGCGTTAGCGCTCAATATCCGCATCCAT 612
QY 773 CTCGAAATCTCTAGTCTGTGATGACCTTTCTTAAAGGCTTAAGACATGCTGAG 832
DB 613 CTCGAAATCTCTAGTCTGTGATGACCTTTCTTAAAGGCTTAAGACATGCTGAG 672
QY 833 CGCCATGACACCGGAGGCTGTGATGACCTTTCTTAAAGGCTTAAGACATGCTGAG 892
DB 673 CGCCATGACACCGGAGGCTGTGATGACCTTTCTTAAAGGCTTAAGACATGCTGAG 732
QY 893 CCCGATGACACCGGAGGCTGTGATGACCTTTCTTAAAGGCTTAAGACATGCTGAG 952
DB 733 CCCGATGACACCGGAGGCTGTGATGACCTTTCTTAAAGGCTTAAGACATGCTGAG 792
QY 953 TATCTGAGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 1012
DB 793 TATCTGAGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 852
QY 1013 AACGTCAGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 1072
DB 853 AACGTCAGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 912
QY 1073 CAGCTGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 1132
DB 913 CAGCTGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 972
QY 1133 GTACGTGACATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 1192
DB 973 GTACGTGACATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 1032
QY 1193 TCAATGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 1252
DB 1033 TCAATGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 1092
QY 1253 TCTAGTTTCAATTTCTCTGCTGCTCTTAAAGAGAGGAGGAGGAGGAGGAGGAGGAG 1312
DB 1093 TCTAGTTTCAATTTCTCTGCTGCTCTTAAAGAGAGGAGGAGGAGGAGGAGGAGGAG 1152
QY 1313 CGTGGGCTCTTTAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 1372

```


Db 1153 CGTGGGCTCTTAGGAGGAGGCTCAAAATAGGAGAGAGCTGAGTCTCCAGAGGGC 1212
Qy 1373 TCTGACAGACAGCGCTGAGGTGATGAGGTGAGGGGCTTCTCAGAGAAACCTGGA 1432
Db 1213 TCTGACAGACAGCGCTGAGGTGATGAGGTGAGGGGCTTCTCAGAGAAACCTGGA 1272
Qy 1433 GCTGTGGGAGAGAGATTGAGAGAGAGAGAAAGCTGCTGCTGCTGAGAGACTTT 1492
Db 1273 GCTGTGGGAGAGAGATTGAGAGAGAGAGAAAGCTGCTGCTGCTGAGAGACTTT 1332
Qy 1493 GAGAGCAATGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 1552
Db 1333 GAGAGCAATGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 1392
Qy 1553 CAGAAATG 1560
Db 1393 CAGAAATG 1400

RESULT 15

US-09-849-869A-30
Sequence 30, Application US/09849869A

GENERAL INFORMATION:
APPLICANT: Anderson, David J.
APPLICANT: Dong, Xinzhang
APPLICANT: Zylka, Mark
APPLICANT: Simon, Melvin
APPLICANT: Han, Sang-kyou
TITLE OF INVENTION: PAIN SIGNALING MOLECULES
FILE REFERENCE: CALTE.004C1
CURRENT APPLICATION NUMBER: US/09/849, 869A
CURRENT FILING DATE: 2001-05-04
PRIOR APPLICATION NUMBER: US 60/222,344
PRIOR FILING DATE: 2000-08-01
PRIOR APPLICATION NUMBER: US 60/202,027
PRIOR FILING DATE: 2000-05-04
PRIOR APPLICATION NUMBER: US 09/704,707
PRIOR FILING DATE: 2000-11-03
PRIOR APPLICATION NUMBER: US 60/285,493
PRIOR FILING DATE: 2001-04-19
NUMBER OF SEQ ID NOS: 115
SOFTWARE: FastSeq for Windows Version 4.0.
SEQ ID NO 30
LENGTH: 1400
TYPE: DNA
ORGANISM: Homo sapiens
FEATURE:
NAME/KEY: CDS
LOCATION: (332)...(1297)
US-09-849-869A-30

Query Match 41.6%; Score 1088; DB 35; Length 1400;
Best Local Similarity 100.0%; Pred. No. 1,4e-225;
Matches 1088; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 473 CAGAGTGGGGTTCTGAGCATGATCAACATCCAGTCTGGGTACAGAACTGAGACC 532
Db 313 CAGAGTGGGGTTCTGAGCATGATCAACATCCAGTCTGGGTACAGAACTGAGACC 372
Qy 533 AATCAACGAGAGTGAAGAGACTCTTGTCTACAGACAGACCTTGAAGCTTCAAGGGGCTGAC 592
Db 373 AATCAACGAGAGTGAAGAGACTCTTGTCTACAGACAGACCTTGAAGCTTCAAGGGGCTGAC 432
Qy 593 GTGATGCTTCTCTTGTGCGGCTGACAGAGAAAGCGGTTGTGCTGCTGCTGCTGCTGCTG 652
Db 433 GTGATGCTTCTCTTGTGCGGCTGACAGAGAAAGCGGTTGTGCTGCTGCTGCTGCTGCTG 492
Qy 653 CCGATGCGAGAGAAAGCTGTCTCATCTATACATCTCAACCTGAGCGGCGGAGCTTCTCT 712
Db 493 CCGATGCGAGAGAAAGCTGTCTCATCTATACATCTCAACCTGAGCGGCGGAGCTTCTCT 552
Qy 713 CTTCCTTAGCGGACCATATATATGTCGCGTTACGCTTCAATATATCCGCATCCCAT 772
Db 1393 CAGAAATG 1400

Db 553 CTTCCTTAGCGGACCATATATATGTCGCGTTACGCTTCAATATATCCGCATCCCAT 612
Qy 773 CTTCAAATCTGAGTCTGAGTGAATGACTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCT 832
Db 613 CTTCAAATCTGAGTCTGAGTGAATGACTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCT 672
Qy 833 CGCATGAGACCGAGGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 892
Db 673 CGCATGAGACCGAGGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 732
Qy 893 CCCAGATACCTGTATGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 952
Db 733 CCCAGATACCTGTATGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 792
Qy 953 TATCTGAGAGATGATGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 1012
Db 793 TATCTGAGAGATGATGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 852
Qy 1013 AACGTGAGATTTCAATTAATCCGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 1072
Db 853 AACGTGAGATTTCAATTAATCCGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 912
Qy 1073 CAGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 1132
Db 913 CAGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 972
Qy 1133 GTACGTGACATCTCTCTCAAGTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 1192
Db 973 GTACGTGACATCTCTCTCAAGTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 1032
Qy 1193 TCAAGTGGGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 1252
Db 1033 TCAAGTGGGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 1092
Qy 1253 TCTAGTTTCAATTTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 1312
Db 1093 TCTAGTTTCAATTTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 1152
Qy 1313 CGTGGGCTCTTTAGGAGAGGCTCAAAATAGGAGAGAGCTGAGCTGCTGCTGCTGCTGCTGCT 1372
Db 1153 CGTGGGCTCTTTAGGAGAGGCTCAAAATAGGAGAGAGCTGAGCTGCTGCTGCTGCTGCTGCT 1212
Qy 1373 TCTGACAGACAGCGCTGAGGTGATGAGGTGAGGGTGTGCTTCTCAGAGAAACCTGGA 1432
Db 1213 TCTGACAGACAGCGCTGAGGTGATGAGGTGAGGGTGTGCTTCTCAGAGAAACCTGGA 1272
Qy 1433 GCTGTGGGAGAGAGATTGAGAGAGAGAGAAAGCTTGTGCTGCTGCTGCTGCTGCTGCTG 1492
Db 1273 GCTGTGGGAGAGAGATTGAGAGAGAGAGAAAGCTTGTGCTGCTGCTGCTGCTGCTGCTG 1332
Qy 1493 GAGAGCAATGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 1552
Db 1333 GAGAGCAATGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 1392
Qy 1553 CAGAAATG 1560
Db 1393 CAGAAATG 1400

Search completed: October 30, 2004, 03:30:29
Job time : 9593 secs

This Page Blank (uspto)

GenCore version 5.1.6
Copyright (c) 1993 - 2004 CompuGen Ltd.

OM nucleic - nucleic search, using SW model

Run on: October 29, 2004, 19:33:38 ; Search time 97 Seconds
(without alignments)
5942.204 Million cell updates/sec

Title: US-09-867-570-1
Perfect score: 2618
Sequence: 1 aacactgcgcgaattcgs.....gaaaaaaaaaaaaaaaaa 2618

Scoring table: OLIGO_NUC
Gapop 60.0 , Gapext 60.0

Searched: 680018 seqs, 110082838 residues

Word size : 20

Total number of hits satisfying chosen parameters: 3851

Minimum DB seq length: 0
Maximum DB seq length: 200000000

Post-processing: Listing first 45 summaries

Database :

Pending Patents NA New:*
1: /cgn2_6/ptodata/2/pna/US06_NEW_COMB.seq:*
2: /cgn2_6/ptodata/2/pna/US07_NEW_COMB.seq:*
3: /cgn2_6/ptodata/2/pna/US08_NEW_COMB.seq:*
4: /cgn2_6/ptodata/2/pna/US09_NEW_COMB.seq:*
5: /cgn2_6/ptodata/2/pna/US10_NEW_COMB.seq:*
6: /cgn2_6/ptodata/2/pna/US11_NEW_COMB.seq:*
7: /cgn2_6/ptodata/2/pna/US60_NEW_COMB.seq:*
8: /cgn2_6/ptodata/2/pna/US60_NEW_COMB.seq:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	1088	41.6	1400	US-10-957-135-30	Sequence 30, Appl
2	120	4.6	2040	US-10-957-135-15	Sequence 15, Appl
3	97	3.7	1604	US-10-957-135-32	Sequence 32, Appl
4	46	1.8	1300	US-10-957-135-17	Sequence 17, Appl
5	37	1.4	1979	US-10-957-135-62	Sequence 62, Appl
6	37	1.4	2110	US-10-957-135-58	Sequence 58, Appl
7	37	1.4	2401	US-10-957-135-56	Sequence 56, Appl
8	34	1.3	2093	US-10-957-135-54	Sequence 54, Appl
9	34	1.3	2758	US-10-957-135-72	Sequence 72, Appl
10	33	1.3	1519	US-10-957-135-52	Sequence 52, Appl
11	33	1.3	1738	US-10-957-135-74	Sequence 74, Appl
12	31	1.2	1045	US-10-957-135-24	Sequence 24, Appl
13	31	1.2	1088	US-10-957-135-1	Sequence 1, Appl
14	28	1.1	1010	US-10-220-366A-9630	Sequence 9630, Ap
15	26	1.0	1485	US-10-957-135-64	Sequence 64, Appl
16	25	1.0	395	US-10-220-366A-7081	Sequence 7081, Ap
17	25	1.0	573	US-10-956-160-6514	Sequence 6514, Ap
18	25	1.0	573	US-10-956-160-21621	Sequence 21621, Ap
19	25	1.0	578	US-10-956-160-7449	Sequence 7449, Ap
20	25	1.0	578	US-10-956-160-217556	Sequence 217556, Ap
21	25	1.0	586	US-10-956-160-1226	Sequence 1226, Ap
22	25	1.0	586	US-10-956-160-211333	Sequence 211333, Ap
23	25	1.0	980	US-10-957-135-26	Sequence 26, Appl
24	25	1.0	1024	US-10-957-135-22	Sequence 22, Appl
25	25	1.0	2853	US-10-957-135-11	Sequence 11, Appl

26	25	1.0	4293	US-10-811-028A-2	Sequence 2, Appl
27	25	1.0	4332	US-10-811-028A-1	Sequence 1, Appl
28	24	0.9	368	US-10-220-366A-5762	Sequence 5762, Ap
29	24	0.9	574	US-10-956-160-10950	Sequence 10950, A
30	24	0.9	574	US-10-956-160-221057	Sequence 221057, A
31	24	0.9	697	US-10-956-157-8066	Sequence 8066, Ap
32	24	0.9	90100	PCT-US03-40884-1	Sequence 1, Appl
33	24	0.9	152	US-10-956-160-12217	Sequence 12217, A
34	23	0.9	201	US-10-220-366A-3751	Sequence 3751, Ap
35	23	0.9	240	US-10-956-160-11880	Sequence 11880, A
36	23	0.9	240	US-10-956-160-221987	Sequence 221987, A
37	23	0.9	293	US-10-220-366A-10832	Sequence 10832, A
38	23	0.9	385	US-10-857-294-1	Sequence 1400, Ap
39	23	0.9	394	US-10-220-366A-1400	Sequence 372, Ap
40	23	0.9	398	US-10-220-366A-2372	Sequence 7410, Ap
41	23	0.9	400	US-10-220-366A-7410	Sequence 5577, Ap
42	23	0.9	403	US-10-220-366A-5577	Sequence 5873, Ap
43	23	0.9	410	US-10-956-160-5873	Sequence 215980,
44	23	0.9	410	US-10-956-160-215980	
45	23	0.9	410	US-10-956-160-215980	

ALIGNMENTS

RESULT 1
US-10-957-135-30
Sequence 30, Application US/10957135
GENERAL INFORMATION:
APPLICANT: Anderson, David J.
APPLICANT: Dong, Xinzhang
APPLICANT: Zylka, Mark
APPLICANT: Simon, Melvin
APPLICANT: Han, Sang-Kyou
TITLE OF INVENTION: PAIN SIGNALING MOLECULES
FILE REFERENCE: CALTE, 4C1CP1C1
CURRENT APPLICATION NUMBER: US/10/957,135
CURRENT FILING DATE: 2004-09-30
PRIOR APPLICATION NUMBER: US 60/222,344
PRIOR FILING DATE: 2000-08-01
PRIOR APPLICATION NUMBER: US 60/202,027
PRIOR FILING DATE: 2000-05-04
PRIOR APPLICATION NUMBER: US 09/704,707
PRIOR FILING DATE: 2000-11-03
PRIOR APPLICATION NUMBER: US 60/285,493
PRIOR FILING DATE: 2001-04-19
PRIOR APPLICATION NUMBER: US 09/849,869
PRIOR FILING DATE: 2001-05-04
PRIOR APPLICATION NUMBER: US 10/183,116
PRIOR FILING DATE: 2002-06-26
NUMBER OF SEQ ID NOS: 109
SOFTWARE: FastSeq for Windows Version 4.0
SEQ ID NO 30
LENGTH: 1400
TYPE: DNA
ORGANISM: Homo sapiens
FEATURES:
NAME/KEY: CDS
LOCATION: (332)...(1297)
US-10-957-135-30

Query Match 41.6%; Score 1088; DB 6; Length 1400;
Best Local Similarity 100.0%; Pred. No. 0;
Matches 1088; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 473 CAGACTGGGTTCTGAGCATGATTCACCATCCGATCTTGCGGTACAGAACTGACACC 532
DB 313 CAGACTGGGTTCTGAGCATGATTCACCATCCGATCTTGCGGTACAGAACTGACACC 372
QY 533 AATCAACGAGCGTGAAGAGACTCTCTGTAAGAGAGACCTTGAAGCTTCAACGGGCTGAC 592
DB 373 AATCAACGAGCGTGAAGAGACTCTCTGTAAGAGAGACCTTGAAGCTTCAACGGGCTGAC 432

```
QY 593 GTGATGCTTTCCCTTGTGCGGCTGACAGAAAACCGGTTGCTCTGCGCTCTGCGGCTG 652
| | | | |
Db 433 GTGATGCTTTCCCTTGTGCGGCTGACAGAAAACCGGTTGCTCTGCGCTCTGCGGCTG 492
| | | | |
QY 653 CCGATGCGGAGAAACGCTGTCTCATCTACATCTTCAACCTGCTGCGGCGGATCTTCT 712
| | | | |
Db 493 CCGATGCGGAGAAACGCTGTCTCATCTACATCTTCAACCTGCTGCGGCGGATCTTCT 552
| | | | |
QY 713 CTTCCTTACGCGGACCATTTATATGTTGCGCGGTTAGCGGCTCATCAATATCCGATCCCAT 772
| | | | |
Db 553 CTTCCTTACGCGGACCATTTATATGTTGCGCGGTTAGCGGCTCATCAATATCCGATCCCAT 612
| | | | |
QY 773 CTTCAAATCTCTAGCTCTGTGATGACCTTTCTCTTATATAGGCTTAAGCATGCTGAG 832
| | | | |
Db 613 CTTCAAATCTCTAGCTCTGTGATGACCTTTCTCTTATATAGGCTTAAGCATGCTGAG 672
| | | | |
QY 833 CGCATACAGACCCAGGCGCTGCTTCCATCTGTGAGCCCATCTGTGATCACTGCGCGG 892
| | | | |
Db 673 CGCATACAGACCCAGGCGCTGCTTCCATCTGTGAGCCCATCTGTGATCACTGCGCGG 732
| | | | |
QY 893 CCCCAGATACCTGTATGATGATGATGATGATGATGATGATGATGATGATGATGATGATG 952
| | | | |
Db 733 CCCCAGATACCTGTATGATGATGATGATGATGATGATGATGATGATGATGATGATGATG 792
| | | | |
QY 953 TATCCTGAGTGAATGTTCTGTGACTTCTGTGATGATGATGATGATGATGATGATGATG 1012
| | | | |
Db 793 TATCCTGAGTGAATGTTCTGTGACTTCTGTGATGATGATGATGATGATGATGATGATG 852
| | | | |
QY 1013 AACGTCAATTTCAATTAATGATGATGATGATGATGATGATGATGATGATGATGATGATG 1072
| | | | |
Db 853 AACGTCAATTTCAATTAATGATGATGATGATGATGATGATGATGATGATGATGATGATG 912
| | | | |
QY 1073 CAGCTGTGCTGCTGTGATGATGATGATGATGATGATGATGATGATGATGATGATGATG 1132
| | | | |
Db 913 CAGCTGTGCTGCTGTGATGATGATGATGATGATGATGATGATGATGATGATGATGATG 972
| | | | |
QY 1133 GTACGTGACATCTCTCTCAAGTGTGCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCT 1192
| | | | |
Db 973 GTACGTGACATCTCTCTCAAGTGTGCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCT 1032
| | | | |
QY 1193 TCAATGAGGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 1252
| | | | |
Db 1033 TCAATGAGGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 1092
| | | | |
QY 1253 TCTAGTTCATTTTCTCTGCTGCTCTTAAACAGAGTGCACACCCCATCAATTAATCTTCT 1312
| | | | |
Db 1093 TCTAGTTCATTTTCTCTGCTGCTCTTAAACAGAGTGCACACCCCATCAATTAATCTTCT 1152
| | | | |
QY 1313 CGTGGGCTCTTTTAAAGCAGCGTCAAAATAGGAGAACTGAAAGCTGTTCTCAAGAGGC 1372
| | | | |
Db 1153 CGTGGGCTCTTTTAAAGCAGCGTCAAAATAGGAGAACTGAAAGCTGTTCTCAAGAGGC 1212
| | | | |
QY 1373 TCTGCAAGACACGCTGAGTGTGATGAAAGTGAAGGAGGCTTCTCAAGAAACCTTGGA 1432
| | | | |
Db 1213 TCTGCAAGACACGCTGAGTGTGATGAAAGTGAAGGAGGCTTCTCAAGAAACCTTGGA 1272
| | | | |
QY 1433 GCTGTGCGGAGACAGATTGAGACAGTGAAGAAACCTTGTGCTGTCAAGAGACTT 1492
| | | | |
Db 1273 GCTGTGCGGAGACAGATTGAGACAGTGAAGAAACCTTGTGCTGTCAAGAGACTT 1332
| | | | |
QY 1493 GAGAGCAATGCTGCGCTGCGACCTTGACATTAATATGATGATTTTCTTACCTTGTGCT 1552
| | | | |
Db 1333 GAGAGCAATGCTGCGCTGCGACCTTGACATTAATATGATGATTTTCTTACCTTGTGCT 1392
| | | | |
QY 1553 CAGAAATG 1560
| | | | |
Db 1393 CAGAAATG 1400
| | | | |
```

RESULT 2
US-10-957-135-15
Sequence 15, Application US/10957135
GENERAL INFORMATION:

```
APPLICANT: Anderson, David J.  
APPLICANT: Dong, Xinzong  
APPLICANT: Zylka, Mark  
APPLICANT: Simon, Melvin  
APPLICANT: Han, Sang-kyou  
TITLE OF INVENTION: PAIN SIGNALING MOLECULES  
FILE REFERENCE: CALTE.4C1CPICT  
CURRENT APPLICATION NUMBER: US/10/957.135  
PRIOR FILING DATE: 2004-09-30  
PRIOR APPLICATION NUMBER: US 60/222.344  
PRIOR FILING DATE: 2000-08-01  
PRIOR APPLICATION NUMBER: US 60/202.027  
PRIOR FILING DATE: 2000-05-04  
PRIOR APPLICATION NUMBER: US 09/704.707  
PRIOR FILING DATE: 2000-11-03  
PRIOR APPLICATION NUMBER: US 60/285.493  
PRIOR FILING DATE: 2001-04-19  
PRIOR APPLICATION NUMBER: US 09/849.869  
PRIOR FILING DATE: 2001-05-04  
PRIOR APPLICATION NUMBER: US 10/183.116  
PRIOR FILING DATE: 2002-06-26  
NUMBER OF SEQ ID NOS: 109  
SOFTWARE: FastSeq for Windows Version 4.0  
SEQ ID NO 15  
LENGTH: 2040  
TYPE: DNA  
ORGANISM: Homo sapiens  
FEATURE:  
NAME/KEY: CDS  
LOCATION: (328) ... (1293)  
US-10-957-135-15  
Query Match 4.6%; Score 120; DB 6; Length 2040;  
Best Local Similarity 100.0%; Pred. No. 3.9e-42;  
Matches 120; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
```

```
QY 1263 ATTTCTGCTGCTGCTTCTTAAACAGAGTGCACACCCCATCAATTAATCTTGTGAGGCTCC 1322
| | | | |
Db 1099 ATTTCTGCTGCTGCTTCTTAAACAGAGTGCACACCCCATCAATTAATCTTGTGAGGCTCC 1158
| | | | |
QY 1322 TTAGGAGCGCTCAAAATAGGAGAACTGAAAGCTGTTCTCAAGAGGCTCTGAGAGAC 1382
| | | | |
Db 1159 TTAGGAGCGCTCAAAATAGGAGAACTGAAAGCTGTTCTCAAGAGGCTCTGAGAGAC 1218
| | | | |
RESULT 3  
US-10-957-135-32  
Sequence 32, Application US/10957135  
GENERAL INFORMATION:  
APPLICANT: Anderson, David J.  
APPLICANT: Dong, Xinzong  
APPLICANT: Zylka, Mark  
APPLICANT: Simon, Melvin  
APPLICANT: Han, Sang-kyou  
TITLE OF INVENTION: PAIN SIGNALING MOLECULES  
FILE REFERENCE: CALTE.4C1CPICT  
CURRENT APPLICATION NUMBER: US/10/957.135  
PRIOR FILING DATE: 2004-09-30  
PRIOR APPLICATION NUMBER: US 60/222.344  
PRIOR FILING DATE: 2000-08-01  
PRIOR APPLICATION NUMBER: US 60/202.027  
PRIOR FILING DATE: 2000-05-04  
PRIOR APPLICATION NUMBER: US 09/704.707  
PRIOR FILING DATE: 2000-11-03  
PRIOR APPLICATION NUMBER: US 60/285.493  
PRIOR FILING DATE: 2001-04-19  
PRIOR APPLICATION NUMBER: US 09/849.869  
PRIOR FILING DATE: 2001-05-04  
PRIOR APPLICATION NUMBER: US 10/183.116  
PRIOR FILING DATE: 2002-06-26  
NUMBER OF SEQ ID NOS: 109  
SOFTWARE: FastSeq for Windows Version 4.0  
SEQ ID NO 32
```

LENGTH: 1604
TYPE: DNA
ORGANISM: Homo sapiens
FEATURE:
NAME/KEY: CDS
LOCATION: (433)...(1398)
US-10-957-135-32

Query Match 3.7%; Score 97; DB 6; Length 1604;
Best Local Similarity 100.0%; Pred. No. 2.5e-32;
Matches 97; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1287 AGTGCACACCCATCATTTACTTCTGCGGCTCTTTAGGACGCTCAAAATAGGACAG 1346
Db 1228 AGTGCACACCCATCATTTACTTCTGCGGCTCTTTAGGACGCTCAAAATAGGACAG 1287

Qy 1347 AACCTGAAGCTGTTCTTCAGAGGCTCTGCAGACA 1383
Db 1288 AACCTGAAGCTGTTCTTCAGAGGCTCTGCAGACA 1324

RESULT 4
US-10-957-135-17
Sequence 17, Application US/10957135
GENERAL INFORMATION:

APPLICANT: Anderson, David J.
APPLICANT: Dong, Xinzhong
APPLICANT: Zylka, Mark
APPLICANT: Simon, Melvin
APPLICANT: Han, Sang-kyou
TITLE OF INVENTION: PAIN SIGNALING MOLECULES
FILE REFERENCE: CALTE 4C1CPI1
CURRENT APPLICATION NUMBER: US/10/957,135
CURRENT FILING DATE: 2004-09-30
PRIOR APPLICATION NUMBER: US 60/222,344
PRIOR FILING DATE: 2000-08-01
PRIOR APPLICATION NUMBER: US 60/202,027
PRIOR FILING DATE: 2000-05-04
PRIOR APPLICATION NUMBER: US 09/704,707
PRIOR FILING DATE: 2000-11-03
PRIOR APPLICATION NUMBER: US 60/285,493
PRIOR FILING DATE: 2001-04-19
PRIOR APPLICATION NUMBER: US 09/849,869
PRIOR FILING DATE: 2001-05-04
PRIOR APPLICATION NUMBER: US 10/183,116
PRIOR FILING DATE: 2002-06-26
NUMBER OF SEQ ID NOS: 109
SOFTWARE: FastSeq for Windows Version 4.0

SEQ ID NO 17
LENGTH: 1300
TYPE: DNA
ORGANISM: Homo sapiens
FEATURE:
NAME/KEY: CDS
LOCATION: (171)...(1160)
US-10-957-135-17

Query Match 1.8%; Score 46; DB 6; Length 1300;
Best Local Similarity 100.0%; Pred. No. 1.5e-10;
Matches 46; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1276 CTCTTACAGCAGGCGAAGCCCATTTACTTCTGCGGCTC 1321
Db 976 CTCTTACAGCAGGCGAAGCCCATTTACTTCTGCGGCTC 1021

RESULT 5
US-10-957-135-62
Sequence 62, Application US/10957135
GENERAL INFORMATION:

APPLICANT: Anderson, David J.
APPLICANT: Dong, Xinzhong
APPLICANT: Zylka, Mark

APPLICANT: Simon, Melvin
APPLICANT: Han, Sang-kyou
TITLE OF INVENTION: PAIN SIGNALING MOLECULES
FILE REFERENCE: CALTE 4C1CPI1
CURRENT APPLICATION NUMBER: US/10/957,135
CURRENT FILING DATE: 2004-09-30
PRIOR APPLICATION NUMBER: US 60/222,344
PRIOR FILING DATE: 2000-08-01
PRIOR APPLICATION NUMBER: US 60/202,027
PRIOR FILING DATE: 2000-05-04
PRIOR APPLICATION NUMBER: US 09/704,707
PRIOR FILING DATE: 2000-11-03
PRIOR APPLICATION NUMBER: US 60/285,493
PRIOR FILING DATE: 2001-04-19
PRIOR APPLICATION NUMBER: US 09/849,869
PRIOR FILING DATE: 2001-05-04
PRIOR APPLICATION NUMBER: US 10/183,116
PRIOR FILING DATE: 2002-06-26
NUMBER OF SEQ ID NOS: 109
SOFTWARE: FastSeq for Windows Version 4.0
SEQ ID NO 62
LENGTH: 1979
TYPE: DNA
ORGANISM: Mus musculus
US-10-957-135-62

Query Match 1.4%; Score 37; DB 6; Length 1979;
Best Local Similarity 100.0%; Pred. No. 1e-06;
Matches 37; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1288 GTGCCAACCCATCATTTACTTCTGCGGCTCCTT 1324
Db 961 GTGCCAACCCATCATTTACTTCTGCGGCTCCTT 997

RESULT 6
US-10-957-135-58
Sequence 58, Application US/10957135
GENERAL INFORMATION:

APPLICANT: Anderson, David J.
APPLICANT: Dong, Xinzhong
APPLICANT: Zylka, Mark
APPLICANT: Simon, Melvin
APPLICANT: Han, Sang-kyou
TITLE OF INVENTION: PAIN SIGNALING MOLECULES
FILE REFERENCE: CALTE 4C1CPI1
CURRENT APPLICATION NUMBER: US/10/957,135
CURRENT FILING DATE: 2004-09-30
PRIOR APPLICATION NUMBER: US 60/222,344
PRIOR FILING DATE: 2000-08-01
PRIOR APPLICATION NUMBER: US 60/202,027
PRIOR FILING DATE: 2000-05-04
PRIOR APPLICATION NUMBER: US 09/704,707
PRIOR FILING DATE: 2000-11-03
PRIOR APPLICATION NUMBER: US 60/285,493
PRIOR FILING DATE: 2001-04-19
PRIOR APPLICATION NUMBER: US 09/849,869
PRIOR FILING DATE: 2001-05-04
PRIOR APPLICATION NUMBER: US 10/183,116
PRIOR FILING DATE: 2002-06-26
NUMBER OF SEQ ID NOS: 109
SOFTWARE: FastSeq for Windows Version 4.0

SEQ ID NO 58
LENGTH: 2110
TYPE: DNA
ORGANISM: Mus musculus
US-10-957-135-58

Query Match 1.4%; Score 37; DB 6; Length 2110;
Best Local Similarity 100.0%; Pred. No. 1e-06;
Matches 37; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1288 GTGCCAACCCATCATTTACTTCTGCGGCTCCTT 1324

Db 1087 GTGCCAACCCCATCTATTACTTCTTCGTGGGCTCCTT 1123

RESULT 7

US-10-957-135-56/c
; Sequence 56, Application US/10957135
; GENERAL INFORMATION:
; APPLICANT: Anderson, David J.
; APPLICANT: Dong, Xinzhang
; APPLICANT: Zylka, Mark
; APPLICANT: Simon, Melvin
; APPLICANT: Han, Sang-kyou
; TITLE OF INVENTION: PAIN SIGNALING MOLECULES
; FILE REFERENCE: CALTE.4C1CPICT1
; CURRENT APPLICATION NUMBER: US/10/957,135
; CURRENT FILING DATE: 2004-09-30
; PRIOR APPLICATION NUMBER: US 60/222,344
; PRIOR FILING DATE: 2000-08-01
; PRIOR APPLICATION NUMBER: US 60/202,027
; PRIOR FILING DATE: 2000-05-04
; PRIOR APPLICATION NUMBER: US 09/704,707
; PRIOR FILING DATE: 2000-11-03
; PRIOR APPLICATION NUMBER: US 60/285,493
; PRIOR FILING DATE: 2001-04-19
; PRIOR APPLICATION NUMBER: US 09/849,869
; PRIOR FILING DATE: 2001-05-04
; PRIOR APPLICATION NUMBER: US 10/183,116
; PRIOR FILING DATE: 2002-06-26
; NUMBER OF SEQ ID NOS: 109
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 56
; LENGTH: 2401
; TYPE: DNA
; ORGANISM: Mus musculus
US-10-957-135-56

Query Match 1.4%; Score 37; DB 6; Length 2401;
Best Local Similarity 100.0%; Pred. No. 1e-06;
Matches 37; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1288 GTGCCAACCCCATCTATTACTTCTTCGTGGGCTCCTT 1324
DB 1138 GTGCCAACCCCATCTATTACTTCTTCGTGGGCTCCTT 1102

RESULT 8

US-10-957-135-54
; Sequence 54, Application US/10957135
; GENERAL INFORMATION:
; APPLICANT: Anderson, David J.
; APPLICANT: Dong, Xinzhang
; APPLICANT: Zylka, Mark
; APPLICANT: Simon, Melvin
; APPLICANT: Han, Sang-kyou
; TITLE OF INVENTION: PAIN SIGNALING MOLECULES
; FILE REFERENCE: CALTE.4C1CPICT1
; CURRENT APPLICATION NUMBER: US/10/957,135
; CURRENT FILING DATE: 2004-09-30
; PRIOR APPLICATION NUMBER: US 60/222,344
; PRIOR FILING DATE: 2000-08-01
; PRIOR APPLICATION NUMBER: US 60/202,027
; PRIOR FILING DATE: 2000-05-04
; PRIOR APPLICATION NUMBER: US 09/704,707
; PRIOR FILING DATE: 2000-11-03
; PRIOR APPLICATION NUMBER: US 60/285,493
; PRIOR FILING DATE: 2001-04-19
; PRIOR APPLICATION NUMBER: US 09/849,869
; PRIOR FILING DATE: 2001-05-04
; PRIOR APPLICATION NUMBER: US 10/183,116
; PRIOR FILING DATE: 2002-06-26
; NUMBER OF SEQ ID NOS: 109
; SOFTWARE: FastSeq for Windows Version 4.0

; SEQ ID NO 54
; LENGTH: 2093
; TYPE: DNA
; ORGANISM: Mus musculus
US-10-957-135-54

Query Match 1.3%; Score 34; DB 6; Length 2093;
Best Local Similarity 100.0%; Pred. No. 1.9e-05;
Matches 34; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1288 GTGCCAACCCCATCTATTACTTCTTCGTGGGCTC 1321
DB 1371 GTGCCAACCCCATCTATTACTTCTTCGTGGGCTC 1404

RESULT 9

US-10-957-135-72/c
; Sequence 72, Application US/10957135
; GENERAL INFORMATION:
; APPLICANT: Anderson, David J.
; APPLICANT: Dong, Xinzhang
; APPLICANT: Zylka, Mark
; APPLICANT: Simon, Melvin
; APPLICANT: Han, Sang-kyou
; TITLE OF INVENTION: PAIN SIGNALING MOLECULES
; FILE REFERENCE: CALTE.4C1CPICT1
; CURRENT APPLICATION NUMBER: US/10/957,135
; CURRENT FILING DATE: 2004-09-30
; PRIOR APPLICATION NUMBER: US 60/222,344
; PRIOR FILING DATE: 2000-08-01
; PRIOR APPLICATION NUMBER: US 60/202,027
; PRIOR FILING DATE: 2000-05-04
; PRIOR APPLICATION NUMBER: US 09/704,707
; PRIOR FILING DATE: 2000-11-03
; PRIOR APPLICATION NUMBER: US 60/285,493
; PRIOR FILING DATE: 2001-04-19
; PRIOR APPLICATION NUMBER: US 09/849,869
; PRIOR FILING DATE: 2001-05-04
; PRIOR APPLICATION NUMBER: US 10/183,116
; PRIOR FILING DATE: 2002-06-26
; NUMBER OF SEQ ID NOS: 109
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 72
; LENGTH: 2758
; TYPE: DNA
; ORGANISM: Mus musculus
US-10-957-135-72

Query Match 1.3%; Score 34; DB 6; Length 2758;
Best Local Similarity 100.0%; Pred. No. 1.9e-05;
Matches 34; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1288 GTGCCAACCCCATCTATTACTTCTTCGTGGGCTC 1321
DB 1121 GTGCCAACCCCATCTATTACTTCTTCGTGGGCTC 1088

RESULT 10

US-10-957-135-52
; Sequence 52, Application US/10957135
; GENERAL INFORMATION:
; APPLICANT: Anderson, David J.
; APPLICANT: Dong, Xinzhang
; APPLICANT: Zylka, Mark
; APPLICANT: Simon, Melvin
; APPLICANT: Han, Sang-kyou
; TITLE OF INVENTION: PAIN SIGNALING MOLECULES
; FILE REFERENCE: CALTE.4C1CPICT1
; CURRENT APPLICATION NUMBER: US/10/957,135
; CURRENT FILING DATE: 2004-09-30
; PRIOR APPLICATION NUMBER: US 60/222,344
; PRIOR FILING DATE: 2000-08-01
; PRIOR APPLICATION NUMBER: US 60/202,027

```

; PRIOR FILING DATE: 2000-05-04
; PRIOR APPLICATION NUMBER: US 09/704,707
; PRIOR FILING DATE: 2000-11-03
; PRIOR APPLICATION NUMBER: US 60/285,493
; PRIOR FILING DATE: 2001-04-19
; PRIOR APPLICATION NUMBER: US 09/849,869
; PRIOR FILING DATE: 2001-05-04
; PRIOR APPLICATION NUMBER: US 10/183,116
; PRIOR FILING DATE: 2002-06-26
; NUMBER OF SEQ ID NOS: 109
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 52
; LENGTH: 1519
; TYPE: DNA
; ORGANISM: Mus musculus
US-10-957-135-52

```

```

Query Match 1.3%; Score 33; DB 6; Length 1519;
Best Local Similarity 100.0%; Pred. No. 5.2e-05;
Matches 33; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

```

```

Qy 1289 TGGCAACCCATCATTTACTTCTTCGTGGCTC 1321
Db 854 TGGCAACCCATCATTTACTTCTTCGTGGCTC 886

```

```

RESULT 11
US-10-957-135-74
; Sequence 74, Application US/10957135
; GENERAL INFORMATION:
; APPLICANT: Anderson, David J.
; APPLICANT: Dong, Xinzhong
; APPLICANT: Zylka, Mark
; APPLICANT: Simon, Melvin
; APPLICANT: Han, Sang-kyou
; TITLE OF INVENTION: PAIN SIGNALING MOLECULES
; FILE REFERENCE: CALTE.4C1CPI1
; CURRENT APPLICATION NUMBER: US/10/957,135
; CURRENT FILING DATE: 2004-09-30
; PRIOR APPLICATION NUMBER: US 60/222,344
; PRIOR FILING DATE: 2000-08-01
; PRIOR APPLICATION NUMBER: US 60/202,027
; PRIOR FILING DATE: 2000-05-04
; PRIOR APPLICATION NUMBER: US 09/704,707
; PRIOR FILING DATE: 2000-11-03
; PRIOR APPLICATION NUMBER: US 60/285,493
; PRIOR FILING DATE: 2001-04-19
; PRIOR APPLICATION NUMBER: US 09/849,869
; PRIOR FILING DATE: 2001-05-04
; PRIOR APPLICATION NUMBER: US 10/183,116
; PRIOR FILING DATE: 2002-06-26
; NUMBER OF SEQ ID NOS: 109
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 74
; LENGTH: 1738
; TYPE: DNA
; ORGANISM: Mus musculus
US-10-957-135-74

```

```

Query Match 1.3%; Score 33; DB 6; Length 1738;
Best Local Similarity 100.0%; Pred. No. 5.2e-05;
Matches 33; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

```

```

Qy 1292 CAACCCATCATTTACTTCTTCGTGGCTCCTT 1324
Db 794 CAACCCATCATTTACTTCTTCGTGGCTCCTT 826

```

```

RESULT 12
US-10-957-135-24
; Sequence 24, Application US/10957135
; GENERAL INFORMATION:
; APPLICANT: Anderson, David J.

```

```

; APPLICANT: Dong, Xinzhong
; APPLICANT: Zylka, Mark
; APPLICANT: Simon, Melvin
; APPLICANT: Han, Sang-kyou
; TITLE OF INVENTION: PAIN SIGNALING MOLECULES
; FILE REFERENCE: CALTE.4C1CPI1
; CURRENT APPLICATION NUMBER: US/10/957,135
; CURRENT FILING DATE: 2004-09-30
; PRIOR APPLICATION NUMBER: US 60/222,344
; PRIOR FILING DATE: 2000-08-01
; PRIOR APPLICATION NUMBER: US 60/202,027
; PRIOR FILING DATE: 2000-05-04
; PRIOR APPLICATION NUMBER: US 09/704,707
; PRIOR FILING DATE: 2000-11-03
; PRIOR APPLICATION NUMBER: US 60/285,493
; PRIOR FILING DATE: 2001-04-19
; PRIOR APPLICATION NUMBER: US 09/849,869
; PRIOR FILING DATE: 2001-05-04
; PRIOR APPLICATION NUMBER: US 10/183,116
; PRIOR FILING DATE: 2002-06-26
; NUMBER OF SEQ ID NOS: 109
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 24
; LENGTH: 1045
; TYPE: DNA
; ORGANISM: Mus musculus
; FEATURE:
; NAME/KEY: CDS
; LOCATION: (106)...(1020)
US-10-957-135-24

```

```

Query Match 1.2%; Score 31; DB 6; Length 1045;
Best Local Similarity 100.0%; Pred. No. 0.00037;
Matches 31; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

```

```

Qy 1288 GTGGCAACCCATCATTTACTTCTTCGTGGG 1318
Db 872 GTGGCAACCCATCATTTACTTCTTCGTGGG 902

```

```

RESULT 13
US-10-957-135-1
; Sequence 1, Application US/10957135
; GENERAL INFORMATION:
; APPLICANT: Anderson, David J.
; APPLICANT: Dong, Xinzhong
; APPLICANT: Zylka, Mark
; APPLICANT: Simon, Melvin
; APPLICANT: Han, Sang-kyou
; TITLE OF INVENTION: PAIN SIGNALING MOLECULES
; FILE REFERENCE: CALTE.4C1CPI1
; CURRENT APPLICATION NUMBER: US/10/957,135
; CURRENT FILING DATE: 2004-09-30
; PRIOR APPLICATION NUMBER: US 60/222,344
; PRIOR FILING DATE: 2000-08-01
; PRIOR APPLICATION NUMBER: US 60/202,027
; PRIOR FILING DATE: 2000-05-04
; PRIOR APPLICATION NUMBER: US 09/704,707
; PRIOR FILING DATE: 2000-11-03
; PRIOR APPLICATION NUMBER: US 60/285,493
; PRIOR FILING DATE: 2001-04-19
; PRIOR APPLICATION NUMBER: US 09/849,869
; PRIOR FILING DATE: 2001-05-04
; PRIOR APPLICATION NUMBER: US 10/183,116
; PRIOR FILING DATE: 2002-06-26
; NUMBER OF SEQ ID NOS: 109
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 1
; LENGTH: 1088
; TYPE: DNA
; ORGANISM: Mus Musculus
; FEATURE:
; NAME/KEY: CDS

```

; LOCATION: (115)... (1026)
US-10-957-135-1

Query Match 1.2%; Score 31; DB 6; Length 1088;
Best Local Similarity 100.0%; Pred. No. 0.00037;
Matches 31; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1288 GTGCCAACCCTCATTTACTTCTTGGG 1318
DB 878 GTGCCAACCCTCATTTACTTCTTGGG 908

RESULT 14
US-10-220-366A-9630/c
; Sequence 9630, Application US/10220366A
; GENERAL INFORMATION:

; APPLICANT: HYSEQ, INC
; TITLE OR INVENTION: Novel Nucleic Acids and Polypeptides
; FILE REFERENCE: 21272-042
; CURRENT APPLICATION NUMBER: US/10/220,366A
; PRIOR FILING DATE: 2002-08-28
; PRIOR APPLICATION NUMBER: 09/577,409
; PRIOR FILING DATE: 2000-05-18
; PRIOR APPLICATION NUMBER: 09/515,126
; PRIOR FILING DATE: 2001-02-23
; NUMBER OF SEQ ID NOS: 27802
; SOFTWARE: Custom
; SEQ ID NO 9630
; LENGTH: 1010
; TYPE: DNA
; ORGANISM: Homo sapiens
; FEATURE:
; NAME/KEY: misc feature
; LOCATION: (1)... (1010)
; OTHER INFORMATION: n = a,t,c or g
US-10-220-366A-9630

Query Match 1.1%; Score 28; DB 6; Length 1010;
Best Local Similarity 100.0%; Pred. No. 0.0071;
Matches 28; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 221 ATAAATTTCCAGTCTCAGTATTCTT 248
DB 402 ATAAATTTCCAGTCTCAGTATTCTT 375

RESULT 15
US-10-957-135-64
; Sequence 64, Application US/10957135
; GENERAL INFORMATION:

; APPLICANT: Anderson, David J.
; APPLICANT: Dong, Xinzhang
; APPLICANT: Zylka, Mark
; APPLICANT: Simon, Melvin
; APPLICANT: Han, Sang-kyou
; TITLE OR INVENTION: PAIN SIGNALING MOLECULES
; FILE REFERENCE: CALTE.4C1CP1C1
; CURRENT APPLICATION NUMBER: US/10/957,135
; CURRENT FILING DATE: 2004-09-30
; PRIOR APPLICATION NUMBER: US 60/222,344
; PRIOR FILING DATE: 2000-08-01
; PRIOR APPLICATION NUMBER: US 60/202,027
; PRIOR FILING DATE: 2000-05-04
; PRIOR APPLICATION NUMBER: US 09/704,707
; PRIOR FILING DATE: 2000-11-03
; PRIOR APPLICATION NUMBER: US 60/285,493
; PRIOR FILING DATE: 2001-04-19
; PRIOR APPLICATION NUMBER: US 09/849,869
; PRIOR FILING DATE: 2001-05-04
; PRIOR APPLICATION NUMBER: US 10/183,116
; PRIOR FILING DATE: 2002-06-26
; NUMBER OF SEQ ID NOS: 109
; SOFTWARE: FastSeq for Windows Version 4.0

; SEQ ID NO 64
; LENGTH: 1485
; TYPE: DNA
; ORGANISM: Mus musculus
US-10-957-135-64

Query Match 1.0%; Score 26; DB 6; Length 1485;
Best Local Similarity 100.0%; Pred. No. 0.05;
Matches 26; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1288 GTGCCAACCCTCATTTACTTCTTC 1313
DB 926 GTGCCAACCCTCATTTACTTCTTC 951

Search completed: October 30, 2004, 03:32:11
Job time : 102 secs

GenCore version 5.1.6
Copyright (c) 1993 - 2004 CompuGen Ltd.

OM nucleic - nucleic search, using sw model

Run on: October 29, 2004, 18:51:18 ; Search time 7918 Seconds
(without alignments)
12048.338 Million cell updates/sec

Title: US-09-867-570-1

Perfect score: 2618
Sequence: 1 aacacatgcgcgcgaattcgg.....gaaaaaaaaaaaaaaaaa 2618

Scoring table: OLIGO_NUC
Gapop 60.0 , Gapept 60.0

Searched: 32822875 seqs, 18219865908 residues

Word size : 20

Total number of hits satisfying chosen parameters: 317745

Minimum DB seq length: 0
Maximum DB seq length: 200000000

Post-processing: Listing first 45 summaries

Database :
EST:
1: gb_esc1:*
2: gb_esc2:*
3: gb_esc3:*
4: gb_esc4:*
5: gb_esc5:*
6: gb_esc6:*
7: gb_esc7:*
8: gb_esc8:*
9: gb_esc9:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length DB	ID	Description
1	550	21.0	889	6	CA455045 AGENCOURT
2	530	20.2	688	2	BR439409 HTM1-025F
3	313	12.0	662	1	AV731610 AV731610
4	269	10.3	905	6	CD243595 AGENCOURT
5	120	4.6	764	4	BG198766 RST18035
6	120	4.6	1009	7	CN837763 AGENCOURT
7	115	4.4	871	4	BG210740 RST30287
8	97	3.7	827	7	CN831885 AGENCOURT
9	97	3.7	835	7	CN843633 AGENCOURT
10	97	3.7	863	7	CN835944 AGENCOURT
11	97	3.7	895	7	CN840024 AGENCOURT
12	97	3.7	924	7	CN839634 AGENCOURT
13	97	3.7	953	7	CN839765 AGENCOURT
14	89	3.4	184	4	BG198197 RST14376
15	85	3.2	248	4	BG196705 RST15931
16	82	3.1	515	1	AL705589 DKE2P686K
17	81	3.1	552	8	AO392205 C17B1-E1-
18	69	2.6	825	7	CN843258 AGENCOURT
19	67	2.6	1207	7	CN831816 AGENCOURT
20	57	2.2	924	6	CD051096 AGENCOURT
21	46	1.8	503	1	AF003828 AF003828
22	44	1.7	468	8	B74348 B74348
23	44	1.7	825	7	CN835770 AGENCOURT
24	44	1.7	826	7	CN839711 AGENCOURT

25	44	1.7	835	7	CN834097
26	44	1.7	851	7	CN839827
27	44	1.7	899	7	CN843534
28	43	1.6	419	1	AL118979
29	43	1.6	855	7	CN839609
30	41	1.6	435	8	AO192307
31	41	1.6	811	7	CN831931
32	41	1.6	871	7	CN839125
33	41	1.6	954	7	CN842891
34	40	1.5	620	8	AO696198
35	39	1.5	618	8	AO587244
36	39	1.5	1114	7	CN837998
37	38	1.5	636	9	AG013749
38	38	1.5	692	9	AG013779
39	38	1.5	694	9	AG013746
40	38	1.5	698	9	AG013748
41	38	1.5	787	4	BG208126
42	37	1.4	503	8	AO888076
43	36	1.4	191	7	T59759
44	36	1.4	598	8	AO583508
45	35	1.3	141	7	H49498

ALIGNMENTS

RESULT 1
CA455045
LOCUS
DEFINITION
AGENCOURT_10735784 MABCL Homo sapiens CDNA clone IMAGE:6722551 5', mRNA sequence.
ACCESSION
CA455045
VERSION
CA455045.1 GI:24905369
KEYWORDS
EST.
SOURCE
Homo sapiens (human)
ORGANISM
Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.
REFERENCE
NIH-MGC http://mgc.ncl.nih.gov/
AUTHORS
National Institutes of Health, Mammalian Gene Collection (MGC)
TITLE
Unpublished (1999)
JOURNAL
Contact: Robert Strausberg, Ph.D.
COMMENT
Email: cgabbs-remail.nih.gov
Tissue Procurement: Kristi A. Eglund, Ira Pastan
CDNA Library Preparation: Invitrogen Corp
CDNA Library Arrayed by: The I.M.A.G.E. Consortium (LNL)
DNA Sequencing by: Agencourt Bioscience Corporation
Clone distribution: MGC clone distribution information can be found through the I.M.A.G.E. Consortium/LNL at:
http://image.llnl.gov
Plate: LLM44285 Row: h Column: 07
High quality sequence stop: 743.
Location/Qualifiers
1. .889
/organism="Homo sapiens"
/mol_type="mRNA"
/db_xref="taxon:9606"
/clone="IMAGE:6722551"
/cell_line="ZR-75-1, MCF7, SK-BR-3, MDA-MB-231, HTERT-HME1, LNCaP"
/lab_host="EMDH10B"
/clone_lib="MAPc1"
/note="Vector: pCMV-SPORT6, Site 1: EcoRV, Site 2: Not I; Subcloned with brain, liver, lung, kidney and muscle. Directionally cloned. Priming method: oligo-dT. Average insert size: 1800 bp. Library amplification: 26,000 fold. Kristi A. Eglund, James J. Vincent, Robert Strausberg, Bungkok Lee & Ira Pastan. Discovery of new breast cancer genes encoding membrane and secreted proteins. Manuscript submitted."

FEATURES

source

ORIGIN

351 Guo Shoujing Road, Zhangjiang Hi-Tech Park, Pudong, Shanghai
201203, P. R. China
Tel: 86-21-50801919 (ex.45)
Fax: 86-21-50801922
Email: hanzg@chgc.sh.cn

This clone is available at CHGC in Shanghai.
Location/Qualifiers

FEATURES
source

1..662
/organism="Homo sapiens"
/mol_type="mRNA"
/db_xref="taxon:9606"
/clone="HTFAD08"
/issue_type="Hypothalamus"
/dev_stage="Adult"
/lab_host="SOLR"
/clone_1ib="HTF"
/note="Vector: pBluescript sk(-); Site_1: EcoRI; Site_2: XhoI"

ORIGIN

Query Match 12.0%; Score 313; DB 1; Length 662;
Best Local Similarity 99.4%; Pred. No. 8e-147;
Matches 463; Conservative 0; Mismatches 3; Indels 0; Gaps 0;

QY 1591 ATGTTATCTACCTGACAGTTCAGTTCACCAATGATTTCCCTGATTCGACGATTC 1650
DB 32 ATGTTATCTACCTGACAGTTCAGTTCACCAATGATTTCCCTGATTCGACGATTC 91

QY 1651 AATGTTGATTCCTGATTCATTCACCAATGATTTCCCTGATTCGACGATTC 1710
DB 92 AATGTTGATTCCTGATTCATTCACCAATGATTTCCCTGATTCGACGATTC 151

QY 1711 TTCTCTACGACACTTTTCTGCACTTTCTGATTCATTAAGAGAGTTCCTCCACAC 1770
DB 152 TTCTCTACGACACTTTTCTGCACTTTCTGATTCATTAAGAGAGTTCCTCCACAC 211

QY 1771 CCTAAACCTCTCTTATACCTGTTCTCTACCTGATTCATTAAGAGAGTTCCTCTTA 1830
DB 212 CCTAAACCTCTCTTATACCTGTTCTCTACCTGATTCATTAAGAGAGTTCCTCTTA 271

QY 1831 TTAATCTGTCAGACTATGTTCCCTGAAAATCATGTTCCCTTATGATTCGAGGCACTTA 1890
DB 272 TTAATCTGTCAGACTATGTTCCCTGAAAATCATGTTCCCTTATGATTCGAGGCACTTA 331

QY 1891 CTCGAGTTGGAAGCTCAATCTTATATAGAGTTCCTGCTAATCTTAATTCATTAAT 1950
DB 332 CTCGAGTTGGAAGCTCAATCTTATATAGAGTTCCTGCTAATCTTAATTCATTAAT 391

QY 1951 TCTCAGATATAGCAAAATATAGACTTATAGAGAGATTCCTGCTCATTAAGAGAGTC 2010
DB 392 TCTCAGATATAGCAAAATATAGACTTATAGAGAGATTCCTGCTCATTAAGAGAGTC 451

QY 2011 TTAGAAATGGTTTATGAAATAGCCCTCTCTGTCATTTGTCACA 2056
DB 452 TTAGAAATGGTTTATGAAATAGCCCTCTCTGTCATTTGTCACA 497

RESULT 4
CD243595 905 bp mRNA linear EST 22-MAY-2003
LOCUS AGENCOURT 14121139 NIH MGC 180 Homo sapiens cDNA clone
DEFINITION IMAGE:30383708 5', mRNA sequence.
ACCESSION CD243595
VERSION CD243595.1 GI:31004059
KEYWORDS EST.
SOURCE Homo sapiens (human)
ORGANISM Homo sapiens
Eukaryote; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Homiidae; Homo.
REFERENCE 1 (bases 1 to 905)
AUTHORS NIH-MGC http://mgi.nci.nih.gov/
TITLE National Institutes of Health, Mammalian Gene Collection (MGC)
JOURNAL Unpublished (1999)

COMMENT

Contact: Robert Strausberg, Ph.D.
Email: cgaaps-remail.nih.gov
Tissue Procurement: Dr. Michael Brownstein
CDNA Library Preparation: Invitrogen Corp
CDNA Library Arrayed by: The I.M.A.G.E. Consortium (LNL)
DNA Sequencing by: Agencourt Bioscience Corporation
Clone distribution: MGC clone distribution information can be
found through the I.M.A.G.E. Consortium/LNL at:
http://image.lnl.gov
plate: NDMA47 row: p column: 21
High quality sequence start: 20
High quality sequence stop: 592.
Location/Qualifiers

FEATURES

source

1..905
/organism="Homo sapiens"
/mol_type="mRNA"
/db_xref="taxon:9606"
/clone="IMAGE:30383708"
/lab_host="DH10B-Ton A (T1 and T5 phage resistance)"
/note="Organ: Testis; Vector: pCMV-SPORT6.1; Site_1: NotI;
Site_2: EcoRV (destroyed); Library is oligo-dT primed and
directionally cloned (EcoRV site is destroyed upon
cloning). Average insert size 1.68 kb. Library was
constructed by (Invitrogen). Note: this is a NIH_MGC
Library."

ORIGIN

Query Match 10.3%; Score 269; DB 6; Length 905;
Best Local Similarity 100.0%; Pred. No. 1.8e-124;
Matches 269; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 466 GGGTCATCAACACGCGGTTCTGAGCATGATTCACCAATCCAGTTCGGGTACAGAC 525
DB 364 GGGTCATCAACACGCGGTTCTGAGCATGATTCACCAATCCAGTTCGGGTACAGAC 423

QY 526 TGACCAATCAACGAGCATGAGAGACTCTGCTCAAGAGAGCCCTGAGCTTCAAG 585
DB 424 TGACCAATCAACGAGCATGAGAGACTCTGCTCAAGAGAGCCCTGAGCTTCAAG 483

QY 586 GGGTCATCAACACGCGGTTCTGAGCATGATTCACCAATCCAGTTCGGGTACAGAC 645
DB 484 GGGTCATCAACACGCGGTTCTGAGCATGATTCACCAATCCAGTTCGGGTACAGAC 543

QY 646 TGGGTCGCCGACGAG 705
DB 544 TGGGTCGCCGACGAG 603

QY 706 ACTTCTCTCTCTTACGCGGACATTTATA 734
DB 604 ACTTCTCTCTCTTACGCGGACATTTATA 632

RESULT 5
BG198766 764 bp mRNA linear EST 21-APR-2001
LOCUS RST18035 Atherys RAGE library Homo sapiens cDNA, mRNA sequence.
DEFINITION BG198766
ACCESSION BG198766
VERSION BG198766.1 GI:13720453
KEYWORDS EST.
SOURCE Homo sapiens (human)
ORGANISM Homo sapiens
Eukaryote; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Homiidae; Homo.
REFERENCE 1 (bases 1 to 764)
AUTHORS Harrington J.J., Sherf B., Rundlett S., Jackson P.D., Perry R.,
Cain S., Leventhal C., Thornton M., Ramchandran R.,
Whittington J., Letner L., Costanzo D., McEligott K., Boeger S.,
Mays R., Smith B., Veloso N., Kike A., Hess J., Cochren K., Lo K.,
Offendachner J., Danzig J. and Ducar M.
Creation of genome-wide protein expression libraries using random
activation of gene expression
Nat. Biotechnol. 19 (5), 440-445 (2001)

MEDLINE 21227151
PUBMED 11329013
COMMENT Contact: Scott J. Cain
Athersys, Inc.
3201 Carnegie Ave, Cleveland, OH 44115, USA
Tel: 216 431 9900
Fax: 216 361 9596
Email: scain@atersys.com
High quality sequence stop: 559.
Location/Qualifiers
1..764
/organism="Homo sapiens"
/mol_type="mRNA"
/db_xref="taxon:9606"
/cell_line="HT1080"
/clone_lib="Athersys RAGE Library"
/note="See 'Creation of Genome-wide Protein Expression Libraries using Random Activation of Gene Expression', Nature Biotechnology, in press. Note that even though the cell type indicated is HT1080, since a random activation method was used, these sequence tags are not necessarily expressed in HT1080 under normal circumstances."

ORIGIN
Query Match 4.6%; Score 120; DB 4; Length 764;
Best Local Similarity 100.0%; Pred. No. 7.9e-49;
Matches 120; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1263 ATTTCCGTCGCGCTTTAAGCAGAGTGCCCAACCCCATTTACTTCTTCGTGGGCTCC 1322
|||||
DB 177 ATTTCCGTCGCGCTTTAAGCAGAGTGCCCAACCCCATTTACTTCTTCGTGGGCTCC 118
|||||

QY 1323 TTTAGCAGCGCTCAAAATAGGAGACCTGAAGCTGTCTCCAGAGGGCTCTGCAGGAC 1382
|||||
DB 117 TTTAGGACGCGCTCAAAATAGGAGACCTGAAGCTGTCTCCAGAGGGCTCTGCAGGAC 58
|||||

RESULT 6
CN837763/c 1009 bp mRNA linear EST 02-JUN-2004
LOCUS CN837763
DEFINITION AGENCOURT 15864212 NIH MGC 145 Homo sapiens cDNA clone
IMAGE:30706611 3', mRNA sequence.
ACCESSION CN837763
VERSION CN837763.1 GI:47943418
KEYWORDS EST.
SOURCE Homo sapiens (human)
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominiidae; Homo.
1 (bases 1 to 1009)
NIH-MGC <http://mgc.nci.nih.gov/>.
National Institutes of Health, Mammalian Gene Collection (MGC)
Unpublished (1999)
Contact: Daniela S. Gerhard, Ph.D.
Office of Cancer Genomics
National Cancer Institute / NIH
Bldg. 31 Rm10A07 Bethesda, MD 20892
Email: c9apbs-remail.nih.gov
Tissue Procurement: GPCR Consortium
CDNA Library Preparation: GPCR Consortium
CDNA Library Arrayed By: The I.M.A.G.E. Consortium (LNL)
DNA Sequencing by: Agencourt Bioscience Corporation
Clone distribution: MGC clone distribution information can be
found through the I.M.A.G.E. Consortium/LNL at:
<http://image.llnl.gov>
Plate: IRB14 row: e column: 04
High quality sequence stop: 464.
Location/Qualifiers
1..1009
/organism="Homo sapiens"
/mol_type="mRNA"
/db_xref="taxon:9606"
/clone="IMAGE:30706611"

FEATURES
SOURCE

/tissue_type="mixed"
/lab_host="DH10B"
/clone_lib="NIH MGC 145"
/note="Vector: pCDNA3.1; Site 1: varies by clone; Site 2:
varies by clone; ORFs were PCR-amplified and cloned into
pCDNA3.1 by the GPCR Consortium. Cloning sites vary by
clone and include the following: 5'-EcoRV-XbaI/XhoI-3',
5'-EcoRV-XbaI/NotI-3', EcoRV (TA cloned, non-directional).
For information about which gene each clone represents,
please visit our anonymous ftp site at
ftp://image.llnl.gov/image/rearrayed_plates/IRB1.presv.dat
a Note: this is a NIH_MGC Library."

ORIGIN
Query Match 4.6%; Score 120; DB 7; Length 1009;
Best Local Similarity 100.0%; Pred. No. 7.9e-49;
Matches 120; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1263 ATTTCCGTCGCGCTTTAAGCAGAGTGCCCAACCCCATTTACTTCTTCGTGGGCTCC 1322
|||||
DB 236 ATTTCCGTCGCGCTTTAAGCAGAGTGCCCAACCCCATTTACTTCTTCGTGGGCTCC 177
|||||

QY 1323 TTTAGCAGCGCTCAAAATAGGAGACCTGAAGCTGTCTCCAGAGGGCTCTGCAGGAC 1382
|||||
DB 176 TTTAGGACGCGCTCAAAATAGGAGACCTGAAGCTGTCTCCAGAGGGCTCTGCAGGAC 117
|||||

RESULT 7
BG210740/c 871 bp mRNA linear EST 21-APR-2001
LOCUS BG210740
DEFINITION RST30287 Athersys RAGE Library Homo sapiens cDNA, mRNA sequence.
ACCESSION BG210740
VERSION BG210740.1 GI:13732427
KEYWORDS EST.
SOURCE Homo sapiens (human)
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominiidae; Homo.
1 (bases 1 to 871)
Harrington, J.J., Sherf, B., Rundlett, S., Jackson, P.D., Perry, R.,
Cain, S., Leventhal, C., Thornton, M., Ramchandran, R.,
Whittington, J., Lerner, L., Costanzo, D., McElligott, K., Booser, S.,
Mays, R., Smith, B., Veloso, N., and Ducar, M.
Offenbacher, J., Danzig, J. and Ducar, M.
Creation of genome-wide protein expression libraries using random
activation of gene expression
Nat. Biotechnol. 19 (5), 440-445 (2001)

JOURNAL
MEDLINE 21227151
PUBMED 11329013
COMMENT Contact: Scott J. Cain
Athersys, Inc.
3201 Carnegie Ave, Cleveland, OH 44115, USA
Tel: 216 431 9900
Fax: 216 361 9596
Email: scain@atersys.com
High quality sequence stop: 456.
Location/Qualifiers
1..871
/organism="Homo sapiens"
/mol_type="mRNA"
/db_xref="taxon:9606"
/cell_line="HT1080"
/clone_lib="Athersys RAGE Library"
/note="See 'Creation of Genome-wide Protein Expression Libraries using Random Activation of Gene Expression', Nature Biotechnology, in press. Note that even though the cell type indicated is HT1080, since a random activation method was used, these sequence tags are not necessarily expressed in HT1080 under normal circumstances."

ORIGIN
Query Match 4.4%; Score 115; DB 4; Length 871;
Best Local Similarity 100.0%; Pred. No. 2.7e-46;

Matches	115;	Conservative	0;	Mismatches	0;	Indels	0;	Gaps	0;
Qy	1268	CGTGTCCGCTCTTAA	CAGCAGT	AGCCATCATCTTCTTCGTGGGCTCTTTAG	1327				
Db	158	CTCTTCGCTCTTAA	CAGCAGT	AGCCATCATCTTCTTCGTGGGCTCTTTAG	99				
Qy	1328	GCAGCGTCAAAATTA	TGGCAGAAC	CTGAAGCTGTTCTCCAGAGGGCTCGCAGGAC	1382				
Db	98	GCAGCGTCAAAATTA	TGGCAGAAC	CTGAAGCTGTTCTCCAGAGGGCTCGCAGGAC	44				
RESULT 8	CN831885/c								
LOCUS	CN831885	827 bp	mRNA	linear	EST 02-JUN-2004				
DEFINITION	AGENCOURT 15669850 NIH_MGC 145 Homo sapiens cDNA clone								
ACCESSION	CN831885								
VERSION	CN831885								
KEYWORDS	EST.								
SOURCE	CN831885.1 GI:47935638								
ORGANISM	Homo sapiens (human)								
REFERENCE	Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Primates; Catarrhini; Homiidae; Homo. 1 (bases 1 to 827)								
AUTHORS	NIH-MGC http://mgc.nci.nih.gov/ .								
TITLE	National Institutes of Health, Mammalian Gene Collection (MGC)								
JOURNAL	Unpublished (1999)								
COMMENT	Contact: Daniela S. Gerhard, Ph.D. Office of Cancer Genomics National Cancer Institute / NIH Bldg. 31 Rm10A07 Bethesda, MD 20892 Email: cgasbp-remail.nih.gov Tissue Procurement: GPCR Consortium CDNA Library Preparation: GPCR Consortium								
FEATURES	CDNA Library Arrayed by: The I.M.A.G.E. Consortium (LNL)								
source	DNA sequencing by: Agencourt Bioscience Corporation								
	Clone distribution: MGC clone distribution information can be found through the I.M.A.G.E. Consortium/LNL at: http://image.lnl.gov								
	Plate: IRB14 row: e column: 10								
	High quality sequence stop: 653.								
	1..827								
	/organism="Homo sapiens"								
	/mol_type="RNA"								
	/db_xref="taxon:9606"								
	/IMAGE="IMAGE:30706617"								
	/clone_type="mixed"								
	/tissue_type="mixed"								
	/lab_host="DH10B"								
	/clone_id="NIH_MGC 145"								
	/note="Vector: pCDNA3.1; Site 1: varies by clone; Site 2: varies by clone; ORFs were PCR-amplified and cloned into pCDNA3.1 by the GPCR Consortium. Cloning sites vary by clone and include the following: 5'-ScorV-Xmni/XhoI-3', 5'-ScorV-Xmni/NotI-3', ScorV (TA cloned, non-directional). For information about which gene each clones represents, please visit our anonymous ftp site at ftp://image.lnl.gov/image/rearrayed_plates/IRB1.presv.dat								
	a Note: this is a NIH_MGC Library."								
ORIGIN									
Query Match	3.7%; Score 97; DB 7; Length 827;								
Best Local Similarity	100.0%; Pred. No. 3.8e-37;								
Matches	97; Conservative 0; Mismatches 0; Indels 0; Gaps 0;								
Qy	1287	AGTGCACACCCATCATCTTACTTCTTCGTGGGCTCTTTAGCAGCGTCAAAATAGCG	1346						
Db	237	AGTGCACACCCATCATCTTACTTCTTCGTGGGCTCTTTAGCAGCGTCAAAATAGCG	178						
Qy	1347	AACCTGAAGCTGTCTCCAGAGGGCTCGCAGGACA	1383						
Db	177	AACCTGAAGCTGTCTCCAGAGGGCTCGCAGGACA	141						

LOCUS	DEFINITION	ACCESSION	VERSION	KEYWORDS	ORGANISM	SOURCE
RESULT 9 CN843633/c	835 bp mRNA linear EST 02-JUN-2004	CN843633	15658848 NIH_MGC_145	Human sapiens cDNA clone		
LOCUS	AGENCOURT_15658848 NIH_MGC_145	CN843633	15658848 NIH_MGC_145	Human sapiens cDNA clone		
DEFINITION	IMAGE:30706619 5', mRNA sequence.					
ACCESSION	CN843633					
VERSION	CN843633.1					
KEYWORDS	EST.					
ORGANISM	Homo sapiens (human)					
SOURCE	Homo sapiens					
REFERENCE	Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.					
AUTHORS	1 (bases 1 to 835)					
TITLE	NIH-MGC http://mgs.nci.nih.gov/.					
JOURNAL	National Institutes of Health, Mammalian Gene Collection (MGC)					
COMMENT	Unpublished (1999) Contact: Daniela S. Gerhard, Ph.D. Office of Cancer Genomics National Cancer Institute / NIH Bldg. 31 Rm10A07 Bethesda, MD 20892 Email: gcgabs-remail.nih.gov Tissue Procurement: GPCR Consortium cDNA Library Preparation: GPCR Consortium cDNA Library Arrayed by: The I.M.A.G.E. Consortium (LMNL) DNA Sequencing by: Agencourt Bioscience Corporation Clone distribution: MGC clone distribution information can be found through the I.M.A.G.E. Consortium/LMNL at: http://image.llnl.gov Plate: IRB14 row: e column: 12 High quality sequence stop: 660. Location/Qualifiers 1. 835 /organism="Homo sapiens" /mol_type="mRNA" /db_xref="taxon:9606" /clone="IMAGE:30706619" /tissue_type="mixed" /lab_host="DH10B" /note="Vector: pCDNA3.1, Site 1: varies by clone; Site 2: varies by clone; ORFs were PCR-amplified and cloned into pCDNA3.1 by the GPCR Consortium. Cloning sites vary by clone and include the following: 5'-EcoRV-XmnI/XhoI-3', 5'-EcoRV-XmnI/NotI-3', EcoRV (TA cloned, non-directional). For information about which gene each clone represents, please visit our anonymous ftp site at ftp://image.llnl.gov/image/rearrayed_plates/IRB1.presv.dat a Note: this is a NIH_MGC Library."					
ORIGIN						
Query Match	3.7%; Score 97; DB 7; Length 835;					
Best Local Similarity	100.0%; Pred. No. 3.8e-37; Indels 0; Gaps 0, 0;					
Matches	97; Conservative 0; Mismatches 0;					
Db	1287 AGTGCACACCCACATCTTACTTCTTGCTGGGCTCCTTAGGCAGCGTCAAATAAGCAG 1346					
Qy	237 AGTGCACACCCACATCTTACTTCTTGCTGGGCTCCTTAGGCAGCGTCAAATAAGCAG 178					
Db	1347 AACCTGAAGCTGGTTCCTCAGAGGGCTCTGCAGACA 1383					
Qy	177 AACCTGAAGCTGGTTCCTCAGAGGGCTCTGCAGACA 141					
RESULT 10						
CN835944/c	863 bp mRNA linear EST 02-JUN-2004	CN835944	15864324 NIH_MGC_145	Human sapiens cDNA clone		
LOCUS	AGENCOURT_15864324 NIH_MGC_145	CN835944	15864324 NIH_MGC_145	Human sapiens cDNA clone		
DEFINITION	IMAGE:30706618 3', mRNA sequence.					
ACCESSION	CN835944					
VERSION	CN835944.1					
KEYWORDS	EST.					
SOURCE	Homo sapiens (human)					

ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Homiidae; Homo.
REFERENCE 1 (bases 1 to 863)
AUTHORS NIH-MGC <http://mgc.nci.nih.gov/>.
JOURNAL National Institutes of Health, Mammalian Gene Collection (MGC)
COMMENT Unpublished (1999)
Contact: Daniela S. Gerhard, Ph.D.
Office of Cancer Genomics
National Cancer Institute / NIH
Bldg. 31 Rm10A07 Bethesda, MD 20892
Email: cgapbs-remail.nih.gov
Tissue Procurement: GPCR Consortium
CDNA Library Preparation: GPCR Consortium
CDNA Library Arrayed by: The I.M.A.G.E. Consortium (LNL)
DNA Sequencing by: Agencourt Bioscience Corporation
Clone distribution: MGC clone distribution information can be
found through the I.M.A.G.E. Consortium/LNL at:
<http://image.llnl.gov>
Plate: IRB14 row: e column: 11
High quality sequence stop: 719.
Location/Qualifiers
1. 863
/organism="Homo sapiens"
/mol_type="mRNA"
/db_xref="taxon:9606"
/clone="IMAGE:30706618"
/tissue_type="mixed"
/lab_host="DH10B"
/clone_lib="NIH_MGC_145"
/note="Vector: pCDNA3.1; Site 1: varies by clone; Site 2:
varies by clone; ORFs were PCR-amplified and cloned into
pCDNA3.1 by the GPCR Consortium. Cloning sites vary by
clone and include the following: 5'-EcoRV-XmiI/XhoI-3',
5'-EcoRV-XmiI/NotI-3', EcoRV (TA cloned, non-directional).
For information about which gene each clones represents,
please visit our anonymous ftp site at
ftp://image.llnl.gov/image/rearrayed_plates/IRB1.presv.dat
a Note: this is a NIH_MGC Library."

FEATURES
source

ORIGIN
Query Match 3.7%; Score 97; DB 7; Length 863;
Best Local Similarity 100.0%; Pred. No. 3.8e-37;
Matches 97; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1287 AGTGCACACCCATCTTCTCTGCGGCTCTTAGGACGCTCAATAATGAGCAG 1346
|||||
DB 237 AGTGCACACCCATCTTCTCTGCGGCTCTTAGGACGCTCAATAATGAGCAG 178
|||||
QY 1347 AACCTGAAGCTGTTCTCCAGAGGGCTCTGCAGACA 1383
|||||
DB 177 AACCTGAAGCTGTTCTCCAGAGGGCTCTGCAGACA 141
|||||

RESULT 11
CN840024/c 895 bp mRNA linear EST 02-JUN-2004
LOCUS AGNCOURT 15864276 NIH_MGC_145 Homo sapiens cDNA clone
DEFINITION IMAGE:30706615 3', mRNA sequence.
ACCESSION CN840024
VERSION CN840024.1 GI:47945679
KEYWORDS EST.
SOURCE Homo sapiens (human)
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Homiidae; Homo.
REFERENCE 1 (bases 1 to 895)
AUTHORS NIH-MGC <http://mgc.nci.nih.gov/>.
JOURNAL National Institutes of Health, Mammalian Gene Collection (MGC)
COMMENT Unpublished (1999)
Contact: Daniela S. Gerhard, Ph.D.
Office of Cancer Genomics
National Cancer Institute / NIH

Bldg. 31 Rm10A07 Bethesda, MD 20892
Email: cgapbs-remail.nih.gov
Tissue Procurement: GPCR Consortium
CDNA Library Preparation: GPCR Consortium
CDNA Library Arrayed by: The I.M.A.G.E. Consortium (LNL)
DNA Sequencing by: Agencourt Bioscience Corporation
Clone distribution: MGC clone distribution information can be
found through the I.M.A.G.E. Consortium/LNL at:
<http://image.llnl.gov>
Plate: IRB14 row: e column: 08
High quality sequence stop: 449.
Location/Qualifiers
1. 895
/organism="Homo sapiens"
/mol_type="mRNA"
/db_xref="taxon:9606"
/clone="IMAGE:30706615"
/tissue_type="mixed"
/lab_host="DH10B"
/clone_lib="NIH_MGC_145"
/note="Vector: pCDNA3.1; Site 1: varies by clone; Site 2:
varies by clone; ORFs were PCR-amplified and cloned into
pCDNA3.1 by the GPCR Consortium. Cloning sites vary by
clone and include the following: 5'-EcoRV-XmiI/XhoI-3',
5'-EcoRV-XmiI/NotI-3', EcoRV (TA cloned, non-directional).
For information about which gene each clones represents,
please visit our anonymous ftp site at
ftp://image.llnl.gov/image/rearrayed_plates/IRB1.presv.dat
a Note: this is a NIH_MGC Library."

FEATURES
source

ORIGIN
Query Match 3.7%; Score 97; DB 7; Length 895;
Best Local Similarity 100.0%; Pred. No. 3.8e-37;
Matches 97; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1287 AGTGCACACCCATCTTCTCTGCGGCTCTTAGGACGCTCAATAATGAGCAG 1346
|||||
DB 236 AGTGCACACCCATCTTCTCTGCGGCTCTTAGGACGCTCAATAATGAGCAG 197
|||||
QY 1347 AACCTGAAGCTGTTCTCCAGAGGGCTCTGCAGACA 1383
|||||
DB 196 AACCTGAAGCTGTTCTCCAGAGGGCTCTGCAGACA 160
|||||

RESULT 12
CN839634/c 924 bp mRNA linear EST 02-JUN-2004
LOCUS AGNCOURT 15864260 NIH_MGC_145 Homo sapiens cDNA clone
DEFINITION IMAGE:30706614 3', mRNA sequence.
ACCESSION CN839634
VERSION CN839634.1 GI:47945289
KEYWORDS EST.
SOURCE Homo sapiens (human)
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Homiidae; Homo.
REFERENCE 1 (bases 1 to 924)
AUTHORS NIH-MGC <http://mgc.nci.nih.gov/>.
JOURNAL National Institutes of Health, Mammalian Gene Collection (MGC)
COMMENT Unpublished (1999)
Contact: Daniela S. Gerhard, Ph.D.
Office of Cancer Genomics
National Cancer Institute / NIH
Bldg. 31 Rm10A07 Bethesda, MD 20892
Email: cgapbs-remail.nih.gov
Tissue Procurement: GPCR Consortium
CDNA Library Preparation: GPCR Consortium
CDNA Library Arrayed by: The I.M.A.G.E. Consortium (LNL)
DNA Sequencing by: Agencourt Bioscience Corporation
Clone distribution: MGC clone distribution information can be
found through the I.M.A.G.E. Consortium/LNL at:
<http://image.llnl.gov>
Plate: IRB14 row: e column: 07

FEATURES
source High quality sequence stop: 602.
Location/Qualifiers
1..924

/organism="Homo sapiens"
/mol_type="mRNA"
/db_xref="taxon:9606"
/clone="IMAGE:30706614"
/issue_type="mixed"
/lab_host="DH10B"
/clone_lib="NIH_MGC_145"
/note="Vector: pCDNA3.1; Site 1: varies by clone; Site 2: varies by clone; ORFs were PCR-amplified and cloned into pCDNA3.1 by the GPCR Consortium. Cloning sites vary by clone and include the following: 5'-EcoRV-XmnI/XhoI-3', 5'-EcoRV-XmnI/NotI-3', EcoRV (TA cloned, non-directional). For information about which gene each clones represents, please visit our anonymous ftp site at ftp://image.llnl.gov/image/rearranged_plates/IRBI.presv.dat a Note: this is a NIH_MGC Library."

ORIGIN

Query Match 3.7%; Score 97; DB 7; Length 924;
Best Local Similarity 100.0%; Pred. No. 3.8e-37;
Matches 97; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1287 AGTGCACCCCACTTACTCTTCTGCGGCTCCTTTAGGAGGCTCAAAATAGGCAG 1346
DB 256 AGTGCACCCCACTTACTCTTCTGCGGCTCCTTTAGGAGGCTCAAAATAGGCAG 197
QY 1347 AACCTGAAGCTGTTCTCCAGAGGGCTCTGCAGAGCA 1383
DB 196 AACCTGAAGCTGTTCTCCAGAGGGCTCTGCAGAGCA 160

RESULT 13
LOCUS CN839765 953 bp mRNA linear EST 02-JUN-2004
DEFINITION IMAGE:30706616 3', mRNA sequence.
ACCESSION CN839765
VERSION CN839765.1 GI:47945420
KEYWORDS EST.
SOURCE Homo sapiens (human)
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.
REFERENCE NIH-MGC http://mgc.nci.nih.gov/
AUTHORS National Institutes of Health, Mammalian Gene Collection (MGC)
TITLE Unpublished (1999)
JOURNAL Contact: Daniela S. Gerhard, Ph.D.
COMMENT Office of Cancer Genomics
National Cancer Institute / NIH
Bldg. 31 Rm10A07 Bethesda, MD 20892
Email: c9abbs-r@mail.nih.gov
Tissue Procurement: GPCR Consortium
CDNA library Preparation: GPCR Consortium
CDNA Library Arrayed by: The I.M.A.G.E. Consortium (LNL)
DNA Sequencing by: Agencourt Bioscience Corporation
Clone distribution: MGC clone distribution information can be found through the I.M.A.G.E. Consortium/LNL at:
http://image.llnl.gov
plate: IRB14 row: e column: 09
High quality sequence stop: 578.
Location/Qualifiers
1..953

FEATURES
source
/organism="Homo sapiens"
/mol_type="mRNA"
/db_xref="taxon:9606"
/clone="IMAGE:30706616"
/issue_type="mixed"
/lab_host="DH10B"
/clone_lib="NIH_MGC_145"

/note="Vector: pCDNA3.1; Site 1: varies by clone; Site 2: varies by clone; ORFs were PCR-amplified and cloned into pCDNA3.1 by the GPCR Consortium. Cloning sites vary by clone and include the following: 5'-EcoRV-XmnI/XhoI-3', 5'-EcoRV-XmnI/NotI-3', EcoRV (TA cloned, non-directional). For information about which gene each clones represents, please visit our anonymous ftp site at ftp://image.llnl.gov/image/rearranged_plates/IRBI.presv.dat a Note: this is a NIH_MGC Library."

ORIGIN

Query Match 3.7%; Score 97; DB 7; Length 953;
Best Local Similarity 100.0%; Pred. No. 3.8e-37;
Matches 97; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1287 AGTGCACCCCACTTACTCTTCTGCGGCTCCTTTAGGAGGCTCAAAATAGGCAG 1346
DB 256 AGTGCACCCCACTTACTCTTCTGCGGCTCCTTTAGGAGGCTCAAAATAGGCAG 197
QY 1347 AACCTGAAGCTGTTCTCCAGAGGGCTCTGCAGAGCA 1383
DB 196 AACCTGAAGCTGTTCTCCAGAGGGCTCTGCAGAGCA 160

RESULT 14
LOCUS BG195197 184 bp mRNA linear EST 21-APR-2001
DEFINITION RST14376 Athersys RAGE Library Homo sapiens cDNA, mRNA sequence.
ACCESSION BG195197
VERSION BG195197.1 GI:13716884
KEYWORDS EST.
SOURCE Homo sapiens (human)
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.
REFERENCE Harrington,J.J., Sherf,B., Rundlett,S., Jackson,P.D., Perry,R., Cain,S., Leventhal,C., Thornton,M., Ramachandran,R., Whittington,J., Lerner,L., Costanzo,D., McElligott,K., Bozzer,S., Mays,R., Smith,B., Veloso,N., Kliska,A., Hess,J., Cochran,K., Lo,K., Offenbacher,J., Danzig,J., and Ducar,M.
TITLE Creation of genome-wide protein expression libraries using random activation of gene expression
JOURNAL Nat. Biotechnol. 19 (5), 440-445 (2001)
MEDLINE 21227151
PUBMED 11329013
COMMENT Contact: Scott J. Cain
Athersys, Inc.
3201 Carnegie Ave, Cleveland, OH 44115, USA
Tel: 216 431 9900
Fax: 216 361 9596
Email: scain@athersys.com
High quality sequence stop: 134.
Location/Qualifiers
1..184

FEATURES
source
/organism="Homo sapiens"
/mol_type="mRNA"
/db_xref="taxon:9606"
/cdl_line="HT1080"
/clone_lib="Athersys RAGE Library"
/note="See 'Creation of Genome-wide Protein Expression Libraries using Random Activation of Gene Expression', Nature Biotechnology, in press. Note that even though the cell type indicated is HT1080, since a random activation method was used, these sequence tags are not necessarily expressed in HT1080 under normal circumstances."

ORIGIN

Query Match 3.4%; Score 89; DB 4; Length 184;
Best Local Similarity 100.0%; Pred. No. 4.2e-33;
Matches 89; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1294 ACCCATCACTTACTCTTCTGCGGCTCCTTTAGGAGGCTCAAAATAGGCAGACTGA 1353

Db 145 ACCCCATCATTTACTTCTTCTGCGGCTCTTAGGACGCTCAAAATAGGACGAACCTGA 86
 QY 1354 AGCTGTTTCCAGAGGGCTCTGCAGGAC 1382
 Db 85 AGCTGTTTCCAGAGGGCTCTGCAGGAC 57

RESULT 15

LOCUS BG196705/c 248 bp mRNA linear EST 21-APR-2001
 DEFINITION RST15931 Athersys RAGE Library Homo sapiens cDNA, mRNA sequence.
 ACCESSION BG196705
 VERSION BG196705.1 GI:13718392

KEYWORDS EST.
 SOURCE Homo sapiens (human)

ORGANISM

Homo sapiens
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 Mammalia; Eutheria; Primates; Catarrhini; Homidae; Homo.

REFERENCE

1 (bases 1 to 248)
 Harrington,J.J., Sherf,B., Rundlett,S., Jackson,P.D., Perry,R.,
 Cain,S., Leventhal,C., Thornton,M., Ramchandran,R.,
 Whittington,J., Lerner,L., Costanzo,D., McElligott,K., Booser,S.,
 Mays,R., Smith,E., Veloso,N., Kikka,A., Hess,J., Cothren,K., Lo,K.,
 Offenbacher,J., Danzig,J. and Ducar,M.

TITLE

Creation of genome-wide protein expression libraries using random
 activation of gene expression

Nat. Biotechnol. 19 (5), 440-445 (2001)

JOURNAL
 MEDLINE
 PUBMED

COMMENT

Contact: Scott J. Cain
 Athersys, Inc.
 3201 Carnegie Ave, Cleveland, OH 44115, USA
 Tel: 216 431 9900
 Fax: 216 361 9596

Email: scain@athersys.com

High quality sequence stop: 198.

FEATURES

source

Location/Qualifiers

1..248
 /organism="Homo sapiens"
 /mol_type="mRNA"
 /db_xref="taxon:9606"
 /cell_line="HT1080"
 /clone_lib="Athersys RAGE Library"
 /note="See 'Creation of Genome-wide Protein Expression'
 Libraries using Random Activation of Gene Expression',
 Nature Biotechnology, in press. Note that even though the
 cell type indicated is HT1080, since a random activation
 method was used, these sequence tags are not necessarily
 expressed in HT1080 under normal circumstances."

ORIGIN

Query Match 3.2%; Score 85; DB 4; Length 248;

Best Local Similarity 100.0%; Pred. No. 4.5e-31;

Matches 85; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1263 ATTTTCGTCGCGCTTTAAGCAGAGTGCCACCCCATATTACTTCTTGGGAGCTCC 1322

Db 175 ATTTTCGTCGCGCTTTAAGCAGAGTGCCACCCCATATTACTTCTTGGGAGCTCC 116

QY 1323 TTATGGCAGCGCTCAAAATAGGCAGA 1347

Db 115 TTATGGCAGCGCTCAAAATAGGCAGA 91

Search completed: October 30, 2004, 00:46:40
 Job time : 7932 secs